UNCLASSIFIED

AD 251 327

Reproduced
by the

MED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

Best Available Copy

UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

AIR-269-60-FR-235

CATALOGED BY ASTIA
AS AD No. 251 337

IDENTIFICATION, DEVELOPMENT AND UTILIZATION OF HUMAN TALENTS

Contract No. Nonr-2596(00)

Final Report

June, 1960

48 700

ASTIA
[illegible]
1961
[illegible]
TIPDR

AMERICAN INSTITUTE for RESEARCH
WASHINGTON, D. C.

Best
Available
Copy

PROJECT TALENT MONOGRAPH SERIES

Monograph No. 1

DESIGNING THE STUDY

John C. Flanagan
John T. Dailey
Marion F. Shaycoft
William A. Gorham
David B. Orr
Isadore Goldberg

June 1960

Final Report for

Office of Naval Research Contract Nonr-2596(00)

29 December 1960

ERRATA FOR PROJECT TALENT MONOGRAPH NO. 1

Chap.	Page	Line	
II	1	6	Change 410,000 to 440,000
	8	8	Change 1357 to 1353
	8	11	Change 410,000 to 440,000
	10	19	Insert electrographic between having and pencil
	10	22	Delete electrographic
	10	25, 26	Change ink or any opaque material to an ordinary pencil
III	3	8	Change 1355 to 1353
	5	18	Change sub-group to sample
	7	24	Change school to school's student body
	8	5	Change 6% to 7%
	13	24	Insert and other salient characteristics between scores and of
	24	3,4	Appendix 3C being published separately as a technical memorandum
	28	6,7	Appendix 3D being published separately as a technical memorandum
IV	8	20	Delete parentheses ( )
	8	33	Delete parentheses ( )
	9	3	Delete parentheses ( )
	11	5,6	Change later in this chapter, under the heading, "Test Reliability" to on pages IV-18 and IV-19
	13	16	Remove dash (--) and substitute that was needed
	14	7	Change ambituous to ambiguous
	14	23	Change color to pigment
	Table IV-1		In last column, interchange Visualization in Two Dimensions and Visualization in Three Dimensions
VI	2	22	Change 400,000 to 440,000
	7	39	Change discussed to presented
VI	12	29	Change imposes to impose
	12	30	Delete more heavily
	13	29	Change Numbers to Cards
	25	11	Change 1357 to 1353
	26	8	Insert III-25 after page
VII	1	1-5	Delete first paragraph: In Section I.........will be presented.
	4	17	Delete and
	4	17	Change options showing to which the options are
	5	22	Delete parentheses ( )
	10	1	Change colon (:) after gauntlet to a double colon (::)
	12	16	Change more to most
	29	24	Change Medicine to Health
	29	26	Insert medical and nursing at end of line after for
	42	19	Insert 6 within the parentheses ( )
	43	14	Change should be to were
	44	21	Change meaningful material to simple statements
	54	11	Change initial R to E
	64	18,19	Delete or defensible
	66	22	Delete therefore
	79	2,3	Delete if any
	79	8	Delete D. It doesn't turn.
	85	1	Add after Item 113. (Figure grouping item)
	86	-	At right of Item 116, change Same options as Item 112. to Same options as Item 111.
	91	25	Insert decimals, between fractions, and percentage
	92	11	Change s$\sqrt{\ }$s to 3$\sqrt{\ }$s
	96	7,8	Change all call for short division to involve, and delete parentheses from remainder of sentence

ERRATA FOR PROJECT TALENT MONOGRAPH NO. 1

Page 2

Chap.	Page	Line	
VII	97	3	Change eliminates to reduces
	98	23	Change subject to task
	99	25	Change imperfections to differences
	99	26	Delete with a standard
	101	14	Change THE TEST IS EXPERIMENTAL to THE TEST IS STILL IN THE EXPERIMENTAL STAGE OF DEVELOPMENT.
	Table VII-1		Please number this as page VII-103 Change item 18 from Medicine to Health
	VII-1 References		Item (10), insert Knoell, D.M., & Harris, C.W. Factor Analysis of Spelling Abilities. J. educ. Res., 1952, 46, 95-111. Please number this as page VII-104
VIII Part 1	6	7	Delete Test, in; insert Technique (1921).
	33	29	Change much to many
VIII Part 2	39	2	Insert acquired and between information and retained
	45	8	Change ('56) to (1949)
	49	11	Change (1945) to (1943)
	49	19	Change (1953) to (1952)
	51	8	Change ('55) to (1943)
	53	9	Insert influencing between to and his education
	55	21	Change (1945) to (1943)
VIII Part 3	62	9	Insert survey between a and study of
	72	21	Change 1943 to 1954
	72	22	Change to read (Gropper and Fitzpatrick,

Chap.	Page	Line	
VIII Part 3	79	13, 14	Change to read: (5) hobbies; (6) work experiences; (7) other activities;
	80	26	Add and Evart after Kelly
	83	22	Insert VIII-79, after page
	88	26, 27	Should read: .....level of responsibility in his job, and their mother's work.
	91	12	Insert and financial habits between gain and in the future.
	93	16	Insert physical between the and activity
	104	24	Under column TO: change Well to Quite well (no change)
IX	14	21	Insert not between were and well
X	10	29	First word: change counselor's to counselors; Seventh word: change counselors to counselor's
VI	9	27	Change probably to probable
VII	15	12	Put 2. before Vocabulary
	19	3	Put 3. before Scales
	30	30	Put 4. before Screening
	35	1	Put 5. before Scientific
	62	15	Change present to presents
	80	13	Change 13a to J-1
	80	15	Change 13b to J-2
	83	1	Change 14 to K
	90	10	Change 16a to M-1
	90	12	Change 16b to M-2
	94	6	In Item 122, Option D, π should precede the period.
	94	18	Change 17 to N
	97	14	Change 18 to O
	98	17	Change 19 to P
	99	21	Change 20 to Q
	100	12	Change 21 to R
	101	26	Change 22 to S

DESIGNING THE STUDY

Table of Contents

PREFACE AND ACKNOWLEDGMENT

Steps were being taken to obtain facts about the talents of our youth several years before the Russians launched their first earth satellite. In January 1957, Dr. John C. Flanagan prepared and submitted the first proposal for Project TALENT: An Inventory of the Aptitude Resources Available for the Nation's Needs in Professional, Technical, and Public Service Fields.

A conference was held in the Washington Office of the American Institute for Research on February 27, 1957. At this meeting the proposal was presented to representatives of government agencies and professional groups. At about the same time the proposal was discussed with representatives of private foundations interested in educational research. The National Science Foundation agreed to coordinate plans for government participation in the study.

Dr. Flanagan, in his capacity as Professor of Psychology at the University of Pittsburgh prepared a special proposal to the Cooperative Research Program of the United States Office of Education requesting support for an important aspect of the study dealing with personal background data. The Advisory Committee for the Cooperative Research Program recommended support for this study and the contract was completed in June 1957. Dr. Robert C. Craig, who had been a Research Scientist on the staff of the American Institute for Research

for two years was transferred to the staff of the University of Pittsburgh to serve as Program Director for the study. He began work full time on Project TALENT on July 1, 1957.

On October 10, 1957, representatives of several federal agencies met with the project staff at the offices of the National Science Foundation in Washington to discuss the value of the coordinated research effort, raise technical questions regarding the conduct of the research project, and to make recommendations for further financial support of the project. The conferees felt the project would have important outcomes of definite value to both national and individual objectives. They recommended support for a thorough planning study.

In the spring of 1958, the National Institute of Mental Health, the Office of Naval Research, and the National Science Foundation negotiated supporting contracts with the University of Pittsburgh and the American Institute for Research to supplement the funds provided by the United States Office of Education to carry out the planning phase of the study. In June, Dr. Robert C. Craig resigned to accept a teaching position. Dr. John T. Dailey replaced him as Program Director.

The plans developed by the Advisory Panels and the project staff were submitted to various agencies in the fall of 1958 with a request for support. Dr. Roy M. Hall, Assistant Commissioner for Research in the United States Office of Education, coordinated government planning. In accordance with a new agreement regarding responsibility for fields of major concern in the post-Sputnik period, this coordination had been shifted from the

National Science Foundation to the Office of Education.

In a meeting with Dr. Roy M. Hall at his office in October 1958, it was the consensus of the representatives of the interested agencies that this research was so vital to the new plans and programs of the Office of Education that it seemed most appropriate for that Office to provide major support for the program. On the basis of these recommendations Dr. Flanagan prepared a proposal requesting the additional support needed for the first five years of the project.

This proposal and plans for the study were reviewed and recommended for approval by the Advisory Committee of the Cooperative Research Program in January 1958. At that time the Committee Chairman was Dr. Ralph W. Tyler, Center for Advanced Study in the Behavioral Sciences, Palo Alto, California.

An important step toward the realization of this research program was taken when the United States Commissioner of Education, Lawrence G. Derthick, approved the proposal that the Cooperative Research Program supply the major portion of the support needed.

These funds together with those provided by the National Institute of Mental Health and the Office of Naval Research enabled the staff to proceed in May 1959 with the tryout of experimental forms of the data collection instruments.

Since the inception of Project TALENT, Dr. John C. Flanagan, Professor of Psychology at the University of Pittsburgh and Director of the American Institute for

for Research, has been the Responsible Investigator for the study. Dr. John T. Dailey of the University of Pittsburgh and AIR has been the Program Director since Dr. Craig resigned in June 1958. Other members of the staff who have taken major responsibility for various aspects of the program include: Marion F. Shaycoft, Supervisor for Measurement Studies; William A. Gorham, Supervisor for Special Studies; David B. Orr, Supervisor for School Characteristics Studies; Isadore Goldberg, Supervisor for Manpower and Guidance Studies; and Clinton A. Neyman, Jr., Supervisor for Services.

Four panels of the nation's leading experts were asked to work with the staff in a planning study for the testing phase. They have been active in all phases since. In addition to experts from these panels, other sampling experts worked with in determining the sampling procedures were: Morris Hansen of the Bureau of the Census, Frederick Stephan of Princeton University, William Cochran of Harvard University, and Phillip J. Rulon of the Harvard Graduate School of Education. The overall Chariman of the Advisory Panels during the planning phase was Dr. John H. Fischer, then Superintendent of the Bal-

timore Public Schools. The present Chairman is Dr. Kenneth E. Oberholtzer, Superintendent of the Denver Public Schools.

The following is a list of the Advisory Panels and their members:

Guidance and Counseling Panel

Chairman, Edward Landy, Director, Division of Counseling Services, Newton (Massachusetts) Public Schools
Ralph F. Berdie, Director, Student Counseling Bureau, University of Minnesota
Bruce E. Shear, Chief, Bureau of Guidance, New York State Education Department
John M. Stalnaker, President, National Merit Scholarship Corporation
David V. Tiedeman, Associate Professor of Education, Harvard University
Arthur E. Traxler, Executive Director, Educational Records Bureau
Leona E. Tyler, Professor of Psychology, University of Oregon

Educational Research Panel

Chairman, Robert J. Keller, Director, University High School, University of Minnesota
Reverend O'Neil C. D'Amour, Assistant to the Executive Secretary, National Catholic Education Association
Paul L. Banfield, Headmaster, Landon School (Maryland)
Warren G. Findley, Assistant Superintendent for Pupil Personnel Services, Atlanta Public Schools
Earl J. McGrath, Professor of Education, Teachers College, Columbia University
Paul T. Rankin, Assistant Superintendent of Schools, Detroit Public Schools
James W. Reynolds, Professor of Education, University of Texas
J. Wayne Wrightstone, Director, Bureau of Educational Research, New York City Board of Education

Testing Problems Panel

Chairman, Robert L. Thorndike, Head, Department of Psychological Foundations and Services, Teachers College, Columbia University
Henry Chauncey, President, Educational Testing Service
Wayne H. Holtzman, Associate Director, Hogg Foundation, University of Texas
A. Paul Horst, Professor of Psychology, University of Washington
Lloyd G. Humphreys, Chairman, Department of Psychology, University of Illinois
E. Lowell Kelly, Director, Bureau of Psychological Services, University of Michigan
Joseph Zubin, Principal Research Scientist, Biometrics Research Unit, New York State Department of Mental Hygiene

Manpower and Sociology Panel

Chairman, Samuel A. Stouffer,* Director, Laboratory of Social Relations, Harvard University
E. Franklin Frazier, Head, Department of Sociology, Howard University
Seymour E. Harris, Chairman, Department of Economics, Harvard University
Donald G. Marquis, Massachusetts Institute of Technology
Irving D. Lorge, Executive Officer, Institute of Psychological Research, Teachers College, Columbia University
C. Joseph Nuesse, Dean, School of Social Science, Catholic University of America
Fred L. Strodtbeck, Associate Professor of Sociology, University of Chicago

An Advisory Council of the presidents and executive secretaries of the professional and scientific associations interested in the project were asked to give advice and counsel and to review the suggestions and plans of the four panels. The Advisory Council in 1959-60 included:

Martin Essex, President, American Association of School Administrators
George E. Watson, President, Council of Chief State School Officers
Edgar Fuller, Executive Secretary, Council of Chief State School Officers
Dugald S. Arbuckle, President, American Personnel and Guidance Association
Arthur A. Hitchcock, Executive Secretary, American Personnel and Guidance Association
Paul E. Klopsteg, President, American Association for the Advancement of Science
Dael Wolfle, Administrative Secretary, American Association for the Advancement of Science
Donald O. Hebb, President, American Psychological Association
John G. Darley, Executive Secretary, American Psychological Association
Robert E. Willis, President, National School Boards Association
W. A. Shannon, Executive Director, National School Boards Association
Cliff Robinson, President, National Association of Secondary School Principals
The Rt. Rev. Msgr. T. Leo Keaveny, President, Secondary School Department, National Catholic Education Association
The Rt. Rev. Msgr. F. G. Hochwald, Executive Secretary National Catholic Education Association

---

* Dr. Stouffer was a most effective chairman until his untimely death in the summer of 1960.

Many professional people outside the staff of Project TALENT gave most generously of their time and efforts. Most outstanding were the 90 Regional Coordinators who were the field representatives for the project. Much of the success Project TALENT has enjoyed can be attributed to their efforts in arranging for the cooperation of the schools in administering the tests. They also served as advisors on administrative procedures for the testing phase of the study.

Regional Coordinators for the project are listed by states:

Alabama
Paul R. Givens, Birmingham-Southern College
Herbert Eber, Birmingham
Alaska*
Arizona and Nevada
Richard E. Schutz, Arizona State University, Tempe
Arkansas
Carter Short, University of Arkansas, Fayetteville
Hardy C. Wilcoxon, University of Arkansas, Fayetteville
California
Alex D. Aloia, Loyola University of Los Angeles
John Caffrey, Director of Research, Palo Alto Unified School District
Frederick J. McDonald, Stanford University
Colorado
Anthony C. Tucker, University of Denver
Connecticut
Joseph Raymond Gerberich, University of Connecticut, Storrs
Delaware
Arthur R. DeLong, University of Delaware, Newark
District of Columbia
William A. Gorham, University of Pittsburgh
Florida
Edward Caldwell, Board of Public Instruction, Manatee County, Bradenton
John V. McQuitty, University of Florida, Gainesville
Georgia
Cameron Fincher, Georgia State College for Business Administration, Atlanta
G. H. Fort, Board of Education, Atlanta
Richard H. Kicklighter, University of Florida, Gainesville
R. T. Osborne, University of Georgia, Athens
Hawaii*
Idaho
Elwyn DeLaurier, State Guidance Supervisor, Department of Education, Boise
Illinois
N. L. Gage, University of Illinois, Champaign
Lyman J. Smith, Illinois State Scholarship Commission, Deerfield
Indiana
N. A. Fattu, Indiana University, Bloomington
H. H. Remmers, Purdue University, Lafayette

---

* Handled by Central Staff

Iowa
  Arthur Mittman, University of Iowa, Iowa City
  Gordon J. Rhum, Iowa State Teachers College, Cedar Falls
  Herbert M. Silvey, Iowa State Teachers College, Cedar Falls
Kansas
  Kenneth E. Anderson, University of Kansas, Lawrence
  Charles B. Watkins, Guidance and Personnel Service
    Kansas State Department of Public Instruction, Topeka
Kentucky
  Ernest McDaniel, University of Kentucky, Lexington
Louisiana
  Robert N. Vidulich, Louisiana State University, Baton Rouge
Maine
  David R. Fink, Jr., University of Maine, Orono
Maryland
  Robert C. Lloyd, Baltimore Public Schools
Massachusetts
  Seth Arsenian, Springfield College
  George S. Elias, Springfield College
  Edward Scanlon, Division of Counseling Services,
    Newton Public Schools, West Newton
Michigan
  Claude L. Nemzek, University of Detroit
  Edwin G. Spacie, Central Michigan University
    Mt. Pleasant (now deceased)
  Buford Stefflre, Michigan State University, East Lansing
  Frank B. Womer, University of Michigan, Ann Arbor
Minnesota
  Ralph F. Berdie, Director, Student Counseling Bureau,
    University of Minnesota
  Robert J. Keller, Director, University High School,
    University of Minnesota
Mississippi
  Roscoe A. Boyer, University of Mississippi, University
  Russell W. Levanway, Millsaps College, Jackson
Missouri
  Joseph L. French, University of Missouri, Columbia
  Robert E. Lefton, Psychological Associates, Clayton
Montana
  William A. Garrison, Eastern Montana College
    of Education, Billings
Nebraska
  Warren Baller, University of Nebraska, Lincoln
New Hampshire and Vermont
  Paul McIntire, University of New Hampshire, Durham
New Jersey
  Albert Thompson, Oradell
New Mexico
  Virginia Keehan, Department of Education, Santa Fe
New York
  Warren W. Coxe, Delmar
  S. David Farr, University of Buffalo
  John M. Skalski, Fordham University
  Percival M. Symonds, Professor Emeritus, Teachers College,
    Columbia University (now deceased)
  Clarence M. Williams, University of Rochester

North Carolina
- Roy N. Anderson, North Carolina State College, Raleigh
- Junius A. Davis, University of North Carolina, Greensboro
- Thomas E. Jeffrey, University of North Carolina, Chapel Hill
- William E. Perry, University of North Carolina, Chapel Hill

North Dakota
- Ralph H. Kolstoe, University of North Dakota, Grand Forks
- Robert E. Larson, State Agricultural College, Fargo
- Grant M. Norem, State Teachers College, Minot

Ohio
- Howard B. Lyman, University of Cincinnati
- Walter S. Nosal, John Carroll University, Cleveland
- Ray Wood, Columbus

Oklahoma
- W. R. Brown, University of Oklahoma, Norman

Oregon
- J. Spencer Carlson, University of Oregon, Eugene

Pennsylvania
- L. Kathryn Dice, Department of Public Instruction, Harrisburg
- Roy B. Hackman, Philadelphia
- C Mauritz Lindvall, University of Pittsburgh

Rhode Island
- Frances E. Dunn, Brown University, Providence

South Carolina
- R. L. Kalmbach, Columbia Public Schools, Columbia
- W. C. McCall, University of South Carolina, Columbia
- Donna S. Young, University of South Carolina, Columbia

South Dakota
- V. Gregory Rosemont, Huron College, Huron

Tennessee
- George E. Copple, Vanderbilt University, Nashville
- Louise W. Cureton, Knoxville

Texas
- Robert P. Anderson, Texas Technological College, Lubock
- H. Paul Kelley, University of Texas, Austin
- Saul B. Sells, Texas Christian University, Fort Worth
- Franklin L. Stovall, University of Houston
- David F. Votaw, Sr., San Marcos

Utah
- Hyrum M. Smith, Department of Public Instruction, Salt Lake City

Virginia
- Richard L. Beard, University of Virginia, Charlottesville
- William A. Gorham, University of Pittsburgh
- Donald J. Herrmann, College of William and Mary, Williamsburgh

Washington
- William C. Budd, Western Washington College of Education, Bellingham
- Glen E. Maier, Eastern Washington College of Education, Cheney

West Virginia
- Walter Jarecke, University of West Virginia, Morgantown

Wisconsin
- Elden A. Bond, Milwaukee Public Schools
- Robert C. Craig, Marquette University, Milwaukee
- Ralph H. Tindall, Milwaukee Public Schools

Wyoming
- R. Duane Andrews, Department of Education, Cheyenne

The following individuals reviewed some of the experimental forms of the tests: John B. Carroll, Paul Diederich, Paul Dressel, Max Engelhart, Grace Fivars, Geraldine Spaulding, Machlin Thomas, and Mary Willis. Grace Fivars wrote the first drafts of the Creativity test; Mary Willis wrote the first draft of the Mechanical Reasoning test; and Geraldine Spaulding reviewed the final forms of all the tests. Joyce Brueckel, Bertha Harper, and Nathan Jaspen advised on planning the analysis.

The following individuals proposed the original drafts of the various chapters:

Chapter I: Rationale for the Study
John C. Flanagan

Chapter II: Background and Planning
John T. Dailey

Chapter III: Selecting the Students and the Schools
John T. Dailey and Marion F. Shaycoft

Chapter IV: How the Tests were Constructed
John T. Dailey and Marion F. Shaycoft

Chapter V: How the Tests were Given
William A. Gorham

Chapter VI: Processing and Analyzing the Data
Marion F. Shaycoft

Chapter VII: Background and Description of the Tests
Marion F. Shaycoft and John T. Dailey

Chapter VIII: The Background and Description of the Student Activities Inventory, the Interest Inventory, and the Student Information Blank
David B. Orr and Isadore Goldberg

Chapter IX: The General School Characteristics Questionnaire
David B. Orr

Chapter X: The Guidance Program Questionnaire and the Counselor's Questionnaire
Isadore Goldberg

Chapter XI: The Minds of American Youth
John T. Dailey and William A. Gorham

Chapter XII: What Have We Learned from Project TALENT?
John T. Dailey

This report represents a joint effort. It was produced through the cooperation of the professional and clerical staff for Project TALENT.

CHAPTER I

RATIONALE FOR THE STUDY

Are we making the most of the talents of our youth? Will our children identify, develop and use their aptitudes and abilities? Experts disagree on the nature and extent of available talents. They disagree on how our schools should help each student to develop his talents. And they further disagree on how the individual and the nation can best use the talents developed.

Would raising the quality of our schools in underprivileged areas double or quadruple our supply of top-level scientists, writers, or statesmen? The author of "Wasted Talent" in the report, <u>The Nation's Children</u> published for the 1960 White House Conference on Children and Youth cites studies which he believes demonstrate that superior education is directly responsible for producing high-level talent. This view is supported by the fact that in the Science Talent Search in which winners are selected solely on a national basis without set quotas from individual states, 40 percent of the winners come from New York State which has less than ten percent of the eligible population. New York State's expenditures

for education are among the highest in the country. It is also common knowledge that certain teachers and certain departments in other parts of the country produce a high proportion of the winners in national contests of creative writing, debating, drama, and art.

Many educational experts disagree with the view that only a fraction of the potential talents of our young people are developed. Articles on the value of scholarship aid have usually suggested that there are a relatively small fraction of the very able students who are unable to obtain college training. A report on college scholarships and the higher education of gifted students indicates that most of the several thousand students in the study obtained scholarships and went on to college. Approximately 2000 of the remaining candidates had very high scores but did not win scholarship support of any kind. Nearly ninety percent of these students went on to college in spite of their failure to obtain financial support. It is clear that when we consider only the top one-half of one percent of the nation's high school graduates, nearly all of them go to college. This would indicate that only a very small fraction of the available supply of high-level talent is currently being wasted.

Experts also disagree on the basic nature of human talents. John B. Watson, the founder of behavioristic psychology, stated that given a normal healthy infant he could by varying the environment build into that child any desired pattern of talents. Few psychologists in America still hold this view today. In Russia, however, there is still strong support for this theory.

If children do differ in their ability to learn, is this

ability a single dimension of mind as implied by the Intelligence Quotient from the Stanford-Binet and other tests of general intelligence, or are children's mental abilities best described in terms of several dimensions or aptitudes? During World War II it was clearly demonstrated that the pattern of aptitudes that made a good pilot was different from that which made a good navigator. A series of twenty aptitude tests was given to thousands of cadets before they started training. A follow-up study of their later success in pilot or navigator training indicated which combination of high test scores identified the superior pilot or navigator.

If human talents include a variety of patterns and we are not limited to the single dimension of academic aptitude identified by tests of general intelligence, how can we identify and develop these patterns? How many important patterns of aptitude are there? How can we identify the patterns, which if developed, will make a superior surgeon, high school English teacher, electronics engineer, or auto mechanic?

The identification of a student's unique pattern of aptitudes is the first important step. If he is to realize his highest potential, his talents must be developed. What types of learning experience will be most effective? Many controversies exist regarding what constitutes the best secondary education program. In the past few years there has been a demand from colleges and universities for increased emphasis on the basic skills of reading and writing, and more formal training in mathematics and science. Many scholars, especially those representing the European point of view, say that in America's effort to achieve secondary education for all, we have sacrificed not only scholarship standards, but

the basic fundamental curriculum as well. The critics say that courses in life adjustment, swimming, and driver training are inadequate substitutes for much needed advanced training in the fields of mathematics and physics essential to developing the talents of our able youth.

Other experts on secondary education argue that the schools have a strong responsibility to train our young people for citizenship and to broaden the horizons of the many students who will not go on to colleges and universities.

There is pressure for larger secondary schools in which a greater variety of courses and more homogeneous grouping of talents is possible. Opponents of this view argue for more personal attention in smaller schools and greater opportunities for the development of leadership in both curricular and extracurricular activities.

Reinforcing the view that poorly qualified teachers are instructing our students in the wrong courses in our secondary schools are the efforts by national subject matter specialists to revise drastically the whole course of instruction in many fields of secondary education, including science, mathematics, and the foreign languages.

The many proposals and counter proposals for changing our school programs offered by committees, commissions and individual critics all appear vulnerable on one important criticism. They represent opinions and lack a sound factual basis.

The final, and perhaps most important aspect of the process of increasing the realization of the individual's potential is the use he makes of his unique patterns of aptitudes and abilities after they have been developed. The problems here are centered around such topics as motivation, creativity and

productivity.

In a recent study of the motivation to work the investigators report that factors such as a good salary, a good supervisor, good working conditions, congenial fellow workers, and a good working environment are sufficient to prevent dissatisfaction with work, but they do not provide the positive incentives necessary for strong and persistent efforts. It is proposed that this positive type of motivation to work is provided when the individual realizes that his efforts make a definite contribution toward achieving a goal which he recognizes to be important. Thus, growth and accomplishment become the principal motivators. This concept of self-realization is in opposition to the popular view that motivation and morale are primarily a function of an appreciative supervisor and a pleasant working environment.

The factors influencing creativity have also been the subject of much recent speculation among those concerned with the problem of helping individuals to make best use of their talents. Some argue that creativity is primarily a product of a favorable environment. Others have conducted studies to show that certain types of creativity can be substantially improved through special training courses. There is also a limited amount of evidence suggesting that creativity results from a special aptitude for seeing the application of a familiar idea in an unusual or novel situation.

There has been much concern regarding the productivity of both individuals and groups. Everyone is familiar with examples of the highly talented individual wasting his talents even after they have been fully developed, because he does not make best use of his aptitudes and abilities. We are

also familiar with persons who, in spite of rather mediocre abilities, have forged ahead to make significant contributions.

These issues regarding the talents of American youth--their identification, development, and use--provided the basic framework for the design of Project Talent. The ensuing chapters outline the steps in planning and carrying out the initial phase of a project to provide a sound factual basis to assist parents, teachers and school administrators in helping each individual to realize his greatest potential.

THE ORIGINAL DOCUMENT WAS OF POOR QUALITY. BEST POSSIBLE REPRODUCTION FROM COPY FURNISHED ASTIA.

CHAPTER II

BACKGROUND AND PLANNING

The year was 1959. In a hotel conference room in the nation's capital, a group of men sat quietly making plans for 1984. While the headlines shouted threats of atomic war, atmospheric pollution, social disintegration, and Big Brother, they were calmly perfecting plans for a 25-year study of American youth and their schools. About 410,000 high school students and their schools were to be studied intensively over a 25-year period.

This was to be America's first census of the aptitudes, achievements, and talent potentials of its youth. The students were to be followed up in one year after graduation from high school and then five years, ten years, and twenty years after, in order to compare their test performances and the school's educational and guidance programs with what the students did as adults.

The year was 1939--20 years before. A similar group of men sat around another conference table in New York City discussing a proposal for a similar study. Dr. John C.

Flanagan was presenting a proposal to the Committee on Measurement and Guidance of the American Council on Education to draw a national sample for inventorying the aptitudes and achievements of American youth. While all agreed on the vital importance of such a project, it was reluctantly concluded after much discussion that the project would not be feasible at that time because of the relatively primitive facilities then existing for scoring tests and processing large masses of data.

Within a few months, World War II erupted. During World War II Dr. Flanagan mobilized a large group of psychologists in a research program supporting our U.S. Air Corps. This group in a short period of three or four years made great strides in the techniques of mass testing programs and were responsible for a substantial improvement in the accuracy of selecting pilots and other aircrew members for training. The other Armed Services also had such programs and made much progress in the state of the art of testing and handling of mass data.

Soon after World War II, several members of the Air Corps group began actively to work toward a national census of aptitudes and achievements. Dr. Robert L. Thorndike of Columbia University was assigned responsibility for research toward determining the feasibility of such a census and the best way of doing it. Later, in 1956, a conference on Air Force problems and test standardization was held under the chairmanship of Dr. Lloyd G. Humphreys. This group outlined tentative specifications for a recommended aptitude census.

The work during World War II emphasized the importance

of follow-up studies to prove that the tests used really work in the sense that scores on selected tests are actually demonstrated to predict future important performance. This and similar work improved the effectiveness of follow-up techniques and greatly increased the general level of "know-how" in this field. An excellent example of this is Thorndike's follow-up study of 10,000 men who had taken the battery of 20 aircrew aptitude tests during World War II. He compared their scores with data regarding occupational field entered and salary level achieved twelve years later.

Planning

As a result of a series of meetings during 1957, several government agencies agreed to support a planning study to determine the Talent Census needs from a national point of view and to prepare a specific plan for the program. This planning study was conducted jointly by the American Institute for Research and the University of Pittsburgh, with support from the National Institute of Mental Health, the National Science Foundation, the Office of Naval Research, and the Cooperative Research Program of the U.S. Office of Education.

During this planning phase, the new impetus given educational and scientific studies following the successful satellite experiments of the USSR resulted in a definition of areas of responsibility in these fields among the federal agencies. At this time support of testing programs at the secondary school level was allocated to the U.S. Office of Education. The National Science Foundation therefore suggested, with the concurrence of the Assistant Commissioner of Education for Research, that responsibility for planning and coordinating

this program be transferred to the U.S. Office of Education.

As a result of this and subsequent decisions, "Project Talent" is financed for the most part by funds from the U.S. Office of Education.

Plans developed through the Panels were reviewed by an Advisory Council composed principally of Presidents and Executive Secretaries of the professional associations most interested in the findings from this study. (See preface for details about Panels and Advisory Council.)

Regional Coordinators (part-time Project field representatives) administered the program locally. They arranged for cooperation of schools, and served as advisors on procedures. These local officials, familiar with the educational needs and resources of the local communities, were the key representatives in the important testing phase.

All results from the survey and the studies which follow will be published and will be available to educators and research specialists. Prior to publication, these studies will be reviewed by the Panels of experts and the Advisory Council. Many educational and research groups have expressed keen interest in these studies and will be kept fully informed. A series of reports will be made to educational agencies throughout the country. Reports will go to the cooperating school systems, to schools of education, to state departments of education, and to private educational associations, as well as to the supporting agencies. Results will also be of interest to parents, to industry, to the military services, and to anyone concerned with education and the optimum development and use of human talents, at all levels.

The responsibility for administering the tests of all sorts and collecting the data about the students and their past experiences was placed directly in the hands of the selected schools. The scientific analysis of the original data and the collection of the follow-up data are being accomplished under the direction of the program staff of the University of Pittsburgh with special assistance from the American Institute for Research.

Collecting the Data

In the late spring of 1960, the two-day battery of approximately 30 psychological, educational, and background measures was administered to the students in five percent of the nation's high schools. There were 440,000 students tested in 1353 secondary schools. This will provide a base line of the measurable characteristics of a large representative sample of high school students. Since it is planned to secure national norms for one complete age group (the fifteen-year-olds), the high school students tested had to be supplemented by a sample of the fifteen-year-olds not enrolled in any high school. Therefore in one-tenth of the school districts in the sample, not only the students in the selected school but also all other fifteen-year-olds residing in the district were tested. This category includes fifteen-year-olds who are still in elementary school, those who have dropped out of school without graduating, and high school graduates. The student response was gratifyingly favorable. A number of news stories reported that attendance was noticeably higher on the days of testing.

The original data, including the individual item responses on life history items, activities, and interest

tests, are being preserved to compare with later samples of important behavioral measures during the follow-up study.

The Tests

The two-day battery of tests was designed to be representative of the types of aptitude, achievement, interest, and personality tests that have been proven useful and important in selection, measurement, and guidance of young people at the secondary school level. The test battery includes questionnaires to obtain important background and life history information regarding the previous experiences of the students, their neighborhood, family, community, etc. There are also measures of their motivations, levels of vocational and economic aspiration, value systems, and personal plans. The tests were designed to measure the functions covered by most of the standardized tests now widely used with high school students.

All tests used in the project were prepared especially for this research, following detailed specifications developed by the staff and Advisory Panel after a year's intensive study. No test is an exact parallel of any existing commercial test. The tests will be withheld from subsequent publication and will be reserved for use in research and in the standardization of other tests. The test types used were:

1. Information
    - a. Vocabulary
    - b. Literature
    - c. Music
    - d. Social Studies
    - e. Mathematics
    - f. Physical Sciences
    - g. Biological Sciences
    - h. Aeronautics and Space
    - i. Electricity and Electronics
    - j. Mechanics
    - k. Farming
    - l. Home Economics
    - m. Sports
    - n. Other areas of information
2. Memory for Sentences
3. Memory for Words
4. Disguised Words
5. Spelling
6. Capitalization
7. Punctuation
8. English Usage
9. Effective Expression
10. Word Functions in Sentences
11. Reading Comprehension
12. Creativity
13. Mechanical Reasoning
14. Visualization in Two Dimensions

15. Visualization in Three Dimensions
16. Abstract Reasoning
17. Arithmetic Reasoning
18. Mathematics - Introductory
19. Mathematics - Advanced
20. Arithmetic Computation
21. Table Reading
22. Clerical Checking
23. Object Inspection
24. Preferences

The Interest Test includes material similar to that in several of the more popular interest inventories now being administered to high school students. If obtains self-reports of interest in a wide variety of occupational and professional activities. The Student Activities Inventory is also of the self-description form and includes material similar to that in many of the popular personality tests specifically designed for high school students. It will be scored to provide indicators of several aspects of the personality and temperament of the normal high school student.

The Student Information Blank was developed especially for the study to obtain factual life history information about the students. Questions were designed to gather information on such topics as school and family background, socio-cultural indicators, activities and organizations, and plans for after high school.

Health questions were developed to determine the current status of the student's health, as well as to record past health history.

The tests are described in detail in Chapter VII.

The Schools in the Study

It was planned to select the schools by a random

process so they would be representative of all the secondary schools in the United States. Five percent of the high schools (grades 9-12) were to be selected. Data were to be gathered for all of the students in each school selected. The selection took into account such factors as size and type of school. Some of the leading experts on sampling problems in the country were used as consultants in deciding how to select the schools. Of the 1357 schools selected, approximately 955 were senior high schools and the others were the associated junior high schools. This number of schools included approximately 410,000 students.

The Follow-Up Study

Each group of high school students will be followed up approximately one year after the completion of their high school training (or when they normally would graduate if they drop out before graduation). Thus, the 12th grade group will be followed up in the summer of 1961; the 11th grade group in 1962; the 10th grade group in 1963; and the 9th grade group in 1964. The students will be asked to report on the occupations and training courses actually entered, the reasons for their choices, their success and satisfaction with the activity chosen, marriage status, present plans (including marriage plans), changes of residence, and other related facts regarding their education, career, and adjustment. This information will then be compared with the original test and background data by appropriate electronic computing equipment, and studies of the relationships of test scores and of other originally gathered information to the status of individuals will then be made.

At this time there will be many studies to estimate the effects of various treatments (such as the school's educational procedures or the counseling procedures) on later behavior such as going to college, entering various occupations, etc. Yearly reports are to be prepared containing the findings in the analysis of each successive grade group. At the end of the one-year follow-up, valuable new information should be available regarding such important behaviors as entering college, becoming unemployed, leaving the farm, entering the Armed Forces, remaining in high school until graduation, etc.

It is also proposed that there be follow-up studies five years, ten years, and twenty years after graduation from high school. This would provide information regarding college courses, the first year of occupational experience following college graduation, entry into graduate training, early work experiences for persons going into professional training, marriage experience, re-enlistment in the military, and ultimately an estimate of the significance of the individual's personal adjustment, satisfactions, and contributions in relation to this potentiality as revealed in the initial testing program. These follow-up studies will be done by maintaining continuous contact with these individuals and by obtaining information from them at periodic intervals. Every effort is being made to plan the study in such a way as to maximize participant cooperation in terms of returning questionnaires, answering correspondence, etc.

As a collateral type of study, a "follow-back" technique has been developed. With this technique it is possible to do, in effect, follow-up studies in some instances where

the original link of communication with the subject has been broken. An example of this might be to obtain a list of all of the fellows in the American Physical Society of the right age range twenty years after the test administration of the study. There should be a substantial number of these, and five percent of them should be represented in the original group of students. The total list of names would be matched with the list in the study. Name, name of parents, birth date, high school attended, year of graduation, plus other information will be available to use in matching records. This group could then be compared with the other students to discover important differences in test scores and behavior of the two groups while in high school.

Processing the Information in the Study

One of the technological break-throughs making this study practical and feasible at this time has been the development of the optical sensing electronic scoring machines. The original electrical test scoring machines were based on having pencil marks on an answer sheet complete the circuits between two tiny blades in the scoring machine and thus send a pulse over a dial indicating the score. This required the use of a special electrographic pencil and was a relatively slow, expensive, and not completely dependable process. With optical sensing, the subject makes a mark on the answer sheet with ink or any opaque material and the machine senses or receives the result by a photoelectric cell. This is a far more reliable procedure and is many times faster. The electronic scoring machine developed by Dr. E. F. Lindquist at the University of Iowa will score, check, print, and punch cards for 13

tests at a time at the rate of 6000 answer sheets per hour. In later operations a computer converts the scores to a standard scale, and prints the individual's name, standard scores, combined scores, and other desired information.

The Lindquist optical scanning automatic punch was also employed. This reads answer sheets and punches the responses onto IBM cards at a rate of 6000 cards per hour.

Electronic data processing procedures and machines now make it possible to store hundreds of bits of detailed information on each of the 410,000 students on tapes and cards in readily accessible form for the many important types of analyses and follow-up studies.

What Will Be Done With The Data

Each school filled out approximately 50 pages of questionnaire material regarding its educational programs, characteristics, activities, guidance program, and counselors. Nearly 100 individual test scores and several hundred items of information about life history, aims, experiences, plans, aspirations, etc., were obtained for each student. Several hundred items of information are available about each school. All of these test and school measures will eventually be related to a number of target or goal pay-off variables which will be measures of important educational outcomes and life outcomes. A basic purpose of the studies for Project TALENT is to study the interrelationships of each test or school measure and each criterion variable, with all other test and school measures held constant. The initial approach will be through the basic procedure of multiple-regression analysis. Here, for the large numbers of the predictor variables (test and school measures) versus a

single goal criterion measure at a time, partial and multiple correlations will be computed.

Ordinary correlations between a predictor variable and a dependent goal criterion measure are usually difficult to interpret and may not really represent the true specific relationship between a pair of measures. A striking example of this occurred in a study of Naval recruiting some years ago. It was found, for the various recruiting areas of the country, there was a high negative relationship between the proportion of families with television sets and the recruits obtained from the area per thousand population. However, a multiple-regression analysis showed that there was actually a zero relationship between television sets and recruit productivity when other measures are considered and held constant. This was a result of TV sets (at that time - 1950) existing mainly in the industrialized areas of the country. There was a true unique relationship between degree of industrialization and recruit productivity. This remained as a strong relationship when about 40 variables were analyzed simultaneously, whereas the TV relationship vanished.

It is planned that such multiple regression studies will be carried out for groups of 90 to 100 measures at a time against a specific criterion measure. Such measures will include graduating from high school, going to college, grades made in high school, grades made in college, the occupation entered, the salary received after a number of years, whether the student has migrated, how well he likes his occupational choice, the types of courses he takes in college, whether he becomes a scientist, or leader, or delinquent, or something else.

Group Comparison Studies

As a further approach in exploring the relationships in the information that has been accumulated, distributions of scores will be examined for large numbers of special groups such as:

- a. Occupational groups
- b. Parental occupation groups
- c. Regional groups
- d. Unusually successful groups (creative)
- e. Unusually unsuccessful groups (delinquents, psychiatric cases, etc.)
- f. Migrants
- g. Types of school
- h. Groups with specific attitudes toward business, saving, spending
- i. Groups planning military careers

After the first round of these studies has indicated potential clusterings of the variables' interrelationships, more complex and sophisticated matched group comparisons will be made. Here, averages and other statistics will be computed for all measures for matched groups (groups that are alike in N dimensions but are different in dimension N plus 1). Examples:

- a. Students alike in sex, grade, region, socio-economic status, high school grades, and total Information Test score, but different in planning to attend college (or skipping a grade, failing a grade, wanting to become a scientist, ability to write a paragraph, self-concept, basic values, etc.)
- b. Schools alike in type, size, and parental occupation mix, but different in student-

teacher ratio (or per pupil expenditure, average teacher salary, number of books in library, whether in a specific experimental curriculum group, age of principal, whether multiple or single track, percent of teachers who are men, courses offered, etc.)

Expected Results

In this kind of scientific study, it is possible to know in advance some of the major types of information which will be produced.

Some of the important results of this National Aptitude and Ability Census will be:

An Inventory of Human Resources. "Project Talent" will be a tremendous inventory or stocktaking in which we find out the capabilities of our youth.

We also plan to study the relationships between one kind of ability and another, between one type of school course and another, and between personal hobbies and the development of many types of competence.

A Set of Standards for Educational and Psychological Measurement. When reliable measures have been obtained from thousands of persons in a systematically selected sample, it will be possible to provide a more accurate set of benchmarks or standards for test authors to use in standardizing tests so that scores indicate comparable levels of ability.

This may be roughly compared to the basic standards such as the marked bars for measuring length, or the standard weights which the National Bureau of Standards maintains.

A Comprehensive Counseling Guide Indicating the Patterns of Aptitude and Ability Which Are Predictive of Success in Various Careers. In the follow-ups after the national examinations and analysis, students who took the tests will be located and asked to report on educational and vocational experiences. A young girl may have become a secretary, or a housewife, or she may have gone to college. By studying thousands of student aptitude, interest, and ability patterns, and finding out the person's later activities and occupations, we will learn a great deal. This will help students by predicting more precisely what kinds of aptitudes and abilities, what kinds of courses, and what kinds of interests constitute the best basis for various kinds of careers. A machinist needs good mechanical ability, a scientist needs mathematical ability, but counselors and teachers know that many other factors enter into the qualifications for success in a career. Motivation is a necessary ingredient, but the best use of the student's special talents requires that he identify this talent early and obtain the education essential for the full development and effective use of his powers.

It is anticipated that "Project Talent" will make a significant contribution toward meeting this need: better prediction, based on actual follow-ups, of a young person's chances for success in a given field. To some extent, this can be done today, and it has been an enormous incentive to teachers, young people, and parents to be told, for example, "This boy unquestionably has talent in a given direction. If he can continue his education, he may reasonably expect to master his chosen trade or profession."

A Better Understanding of How Young People Choose Their Life Work. Many people follow their family trade or profession. They tend to think that people know quite early what their life work will be. Other people drift into an occupation and they tend to think that everyone else more-or-less drifts into a particular trade, business or profession. Many people feel that they have very little choice.

However, many people today do have a choice, and the diversity of occupations and the need for special training continue to increase. We have begun to learn something about the processes by which a young person decides that he would like to be a teacher, a lawyer, or an apprentice for a trade. This study and others can help us understand at what ages certain lifetime careers tend to be chosen.

Naturally, high school seniors talk more about their future careers than do high school freshmen. Yet there are many indications that even younger children already have interests which are related to the fields in which they later work.

A Better Understanding of the Educational Experiences Which Prepare Students for Their Life Work. American education is noted for its diversity. Only through the analysis of detailed information about students, their educational experiences, and their subsequent successes or failures can we hope to make our educational system as flexible and responsive to the individual needs of its students as it must be if our nation is to continue to develop and prosper.

"Project Talent" has been carefully designed to fill an important national need for facts regarding the

identification, development, and utilization of our human resources. This information is intended as a basis for manpower policies, and as a basic resource for the many individuals responsible for the education of our children.

It is hoped that the study may lead to a better understanding of mobility of adults and children and a better understanding of educational and other early life experiences that are conducive to civic leadership, mental health, etc., in adulthood.

First results of initial studies are expected to become available early in 1961. The results of the first follow-up study should become available about eight to twelve months later.

CHAPTER III

SELECTING THE STUDENTS AND THE SCHOOLS

"One school in the mountains of Eastern Kentucky is reached by a small secondary road bounded on either side by steep hills heavily wooded with pine. Occasionally the woods open sufficiently to reveal a cabin with an out building or two, then the trees close in again, the steep slopes silently inhospitable to attempts to use the land to gain a livelihood. At a sharp fork in the road, one turns to the right as directed and the road dribbles out and becomes nonexistent. A wide shallow brook describes a graceful arc as it sweeps along the foot of the hill directly ahead. Across the brook is a school and accompanying dormitories. The dormitories are necessary, for the isolated and rugged country makes daily trips to and from school impossible. This is Red Bird Settlement School."*

---

*Letter from Regional Coordinator

At the corner of W. 205 St. and Golden Ave. in the Bronx, in New York City, students gather from all over the city to study calculus and other advanced courses in science and mathematics. A high proportion of them will eventually go through graduate school and receive a Ph.D. This is the Bronx High School of Science.

Near a school on the south bank of the Potomac River, Gadsby's Tavern still stands in the "Old Town" section of Alexandria, Virginia, flanked by many historic old homes of the same period. The Tavern, a favorite meeting place of George Washington, is just a few blocks from the George Washington High School.

Off a residential street in Washington, D. C., a driveway sweeps into a broad arc up a hill in front of the main entrance of a modern parochial secondary school. The structure is a modern brick and glass design on the outside which carries through to contemporary functional structure on the inside. Swinging through the glass doors, one is immediately impressed by the quiet and efficient atmosphere. Students move from room to room with a bearing which seems to reflect their pride in the physical plant itself. Adjoining the school is a modern athletic field which is in use just about every afternoon after school. This is the Archbishop Carroll High School.

On about the 17th of November 1959 the superintendents for these schools opened their mail and read a letter inviting them to participate in Project Talent. Almost a thousand other superintendents received this letter at the same time. The schools involved make up about five percent of the secondary schools in all 50 states of the United

States. Roughly one school out of every twenty was invited to participate in Project Talent, the first national survey of aptitudes and achievement.

How the Schools were Selected

How did Red Bird Settlement High School happen to participate in the project? How did the Bronx High School of Science happen to participate? And how did each of the other 1,355 high schools happen to be in the project? Did they volunteer? Or were they invited? And if they were in the project by invitation, who decided what schools were to be invited? And on what basis were the decisions made? And furthermore how can a talent "census" be taken without testing all students in all schools?

The answer to these questions is a rather long one and if it gets a little technical in spots, we hope the readers will bear with us, because the selection of the schools was a cornerstone of the project.

Only rarely is a census study based on information gathered from 100 percent of the individuals in the group. For instance, in a high proportion of the surveys carried out by the Bureau of the Census, the Bureau of Labor Statistics, the Department of Agriculture, and other governmental organizations, only part of the group is studied, sometimes a very small part. In industrial and commercial operations too, decisions as to whether the product meets the standards are usually based on inspection of only a very small proportion of the total production. It would be impractical to grade a carload of wheat or a bale of cotton on any other basis. Similarly, only a portion of the total group of American high school students was included

in Project Talent.

But unless the participants in the study were an appropriate group, one that was sufficiently representative of high school students in the country as a whole, the objectives of the study could not be met; it would be impossible to use the results to draw sound inferences about "the American high school" or "the American high school student". The "national census of talents" would be a miscount, and inferences based on it misleading. Hence, in planning the project, there had to be tremendous stress on using sound methods of determining what schools would have the opportunity to participate. Nothing could be left to chance--and somewhat paradoxically almost everything *had* to be left to chance. This remark will probably seem enigmatic to anyone who is not a statistician with specialized knowledge in the area of sampling methodology. The authors hope that by the time the reader has finished this chapter the remark will no longer seem enigmatic.

The Problem of Sampling

The objectives of Project Talent, and the questions to be answered by it, were discussed in Chapter II. The answers to these questions were to be based on high school students. But which ones? The theoretically ideal solution would have been to test *all* high school students in the United States. But this, obviously, would have been utterly impossible from any practical viewpoint. There are over 10,000,000 secondary school students in public and private high schools. Limitations on available money, manpower, and time meant that it would be out of the question to test more than about one-twentieth of these students at the very

most. The question that then arises is "Which twentieth?". This is the essence of the "sampling problem". In its technical sense the word "sample" means a sub-group chosen from a total group (technically called a "population") for the purpose of getting information on the basis of which inferences can be drawn about the total group.

In the case of Project Talent, the basic "population" about which information is sought consists of the total group of high school students in the United States, in Grades 9 through 12.

In any situation where a sample is to be chosen there are three basic ways in which it might be done; let us call these methods the "method of available cases", the "quota method", and the "random sampling method".

The method of available cases is by far the least satisfactory of the three. No principle underlies it except that of expediency. The method is just what its name implies. The only reason for inclusion in the sub-group is willingness to be in it. The way this method would apply to Project Talent might be as follows. Suppose it had been decided to have the Project Talent sample consist of 1300 of the approximately 26,000 high schools in the United States. The first 1300 schools that volunteered would constitute the sample. This would obviously not be a representative group. Volunteers are seldom representative of non-volunteers; ask any Army man! The mere fact of having volunteered would in itself constitute an important difference. Now suppose that fewer than the requisite 1300 schools volunteered spontaneously. The next step, if the "method of available cases" were still being used, would be to write letters to addition-

al schools inviting them to participate. Names of those additional schools would be taken from any handy source--a mailing list that happened to be on hand, the membership list of some association, or any other source or sources that would provide the requisite number of schools. Some of the invited schools would agree to participate and others would decline. But more could be invited until the total number of participants was deemed sufficiently impressive. No matter that they would not be representative! One of the more spectacular examples of the deficiencies of the "method of available cases" was a poll conducted just prior to the 1936 presidential election. The people polled were chosen from lists of automobile registrations, telephone directories, and other convenient sources, and on the basis of this not-at-all representative sample, a Republican victory was predicted. The voters, however, did not cooperate. There was a Democratic landslide and the magazine which had conducted the poll, the Literary Digest, went out of business soon after.

What are the alternatives? Does the cure lie in specifying that certain important characteristics must appear in the sample in the same proportion as in the "population"? For instance, one might specify that the proportion of public, parochial, and private schools in the sample must be the same as in the "population", not only on an overall basis, but also within each state. Having thus set up "quotas" for public schools from Colorado, parochial schools from Michigan, etc., one would still select schools from any list that happened to be handy--or accept those that volunteered--until the quotas were filled.

This is the "quota method". The number of individuals with a particular characteristic or set of characteristics is specified in advance. The fact that these "quotas" are set in such a way that they are at least roughly proportional to the incidence of the characteristics in the "population" as a whole constitutes a great improvement over the "method of available cases". But some of the major disadvantages are still there. The sample is still largely self-selected or selected on the basis of convenience and availability. While the proportions are right in certain specified characteristics, they can be very badly off in regard to other important characteristics.

The only satisfactory solution known is to make the selection in a random manner. One may use either a simple random sample or some other form of "probability sample"--a "stratified random sample", for instance.

One characteristic of a simple random sample is that every member of the "population" has an exactly equal chance of being in the sample. To obtain a simple random sample, some procedure must be set up for drawing the sample on a strictly random basis, and in drawing the sample all members of the "population" must be treated exactly alike.

As applied to Project Talent this would mean that each school must have an opportunity to be selected for the sample and that the probability of being selected must be the same for every school. A reasonable question to ask is "Would this simple random sampling procedure guarantee a representative sample?". And the answer is "No". The sample would probably tend to be representative but the chances that it would be exactly like the population in regard to any speci-

fied determinable characteristics are small indeed. For example, suppose that 7.934% of the high schools in the United States are in California. Does this mean that 7.934% of the schools in the sample would be in California? No, it does not! Chance doesn't operate that way. Maybe fewer than 6% of the schools in the sample would be in California; maybe over 9% would be. In other words, with a large sample the proportions would probably be fairly close but not exact. One would know after the sample had been selected how good it was in this respect. But what about the characteristics that may be important but cannot be measured or otherwise determined? Will the sample be representative in terms of these imponderables? Consider, for example, the enthusiasm of the teachers. Surely this is a factor which has some bearing on the effectiveness of the school. But it is a factor that cannot be quantified accurately at present and perhaps never will be quantifiable. Would a simple random sample of schools be representative in terms of the average level of enthusiasm among the teachers in the school? The answer is the same as for readily determinable characteristics such as the state in which the school is located. The sample would <u>tend</u> towards representativeness, but would probably not be exactly representative. In other words it would be a <u>good</u> sample--but would it be good enough? As a matter of fact it probably <u>would</u> be good enough to give acceptable results; but it is possible to do better than that. The reader will recall that the "quota method" has one advantage that the simple random sampling method lacks. If the quota method is used and if the quotas are set up properly, it is possible to guarantee that the sample will be representative, at least with respect

to the characteristics in terms of which quotas are specified. For instance, if quotas are specified for schools of various sizes, these quotas can be set in such a way that the representation of various-sized schools in the sample will be proportional to their representation in the "population of schools". This advantage of the quota method can be combined with the advantages of random sampling by adopting a device called "stratified random sampling". This consists in dividing the population (e.g., the schools) into groups which are similar in regard to one or more characteristics (for instance, location, size, type, etc.). The "quota" for each of these homogeneous groups, which in the plural are called "strata" (singular: "stratum"), is a specified proportion of the total number of members of the stratum. Thus far this description will probably seem to the reader to be a straight description of the quota method; and that is exactly what it is--thus far. The difference comes in at the point where the sample is actually chosen.

Within each stratum, a strictly random selection procedure is used. Every school in the stratum is treated on exactly the same basis as every other; chance alone determines which schools within the stratum are chosen to constitute the quota. This makes it possible to <u>guarantee</u> that the sample will be representative in regard to the characteristics specified by the strata. True, it is not feasible to specify all the relevant characteristics. Some are imponderables, not capable of being quantified, or indicated in objective terms (teacher enthusiasm was given as an example of this). Other factors that not only are relevant and important but even are capable of being measured may be overlooked, because in our

present stage of knowledge their importance is still unknown or overlooked. Still other factors, which are measurable and of recognized importance, may nevertheless have to be eliminated as bases for stratification merely because there are practical limits to the number of different characteristics on which the "population" should be stratified. If there are too many strata, the number of members in many of them will be too small, so that their theoretical representation in the sample will be fractional--which is always inconvenient and usually impossible. In view of the fact that there is thus an upper limit on the number of "stratification variables" that can be used, what about the relevant variables that are <u>not</u> used as bases for stratification? Will the sample be representative in regard to them? The answer, as in the case of simple random sampling, is that it will tend to, but that it will not be <u>exactly</u> representative. However, it is likely to be more nearly representative on many variables when stratified on some, than it would be if there had been no stratification. In brief, stratification makes the sample more representative in <u>some</u> respects and less representative in <u>none</u>--a fine example of having one's cake and eating it too. In view of this admirable feature, it should come as no surprise to the reader that stratified random sampling was the method chosen for selecting the Project Talent sample. The characteristics on which the stratification was based are discussed later in this chapter.

But before we discuss how we stratified what we sampled, it is necessary to explain just what it was that we <u>did</u> sample. Or to put it more technically, what kind of "sampling unit" was used? There were several possibilities. The

primary sampling unit might have been the student, the school, a specific grade in a specific school, the school system (which in most cases would be the same as the city, town, or township), the county, or even the state. If the first of these alternatives, the individual student, were chosen as the sampling unit, it would mean that each student would be chosen at random individually to be in the sample. The fact that any particular student had been chosen would have no bearing whatever on whether any of his classmates or schoolmates, or any fellow-residents of his town, county, or state were also chosen. At the other extreme, if the state were the primary sampling unit, it would mean that perhaps one or two or more states would be chosen at random from among the fifty, and that then only students in the selected states would be tested. This would not be a good way of doing the sampling for Project Talent. There are enormous differences among states in regard to educational standards and many other highly relevant factors. A small handful of states could hardly be expected to represent the country as a whole very well.

On strictly theoretical grounds the ideal procedure for Project Talent might have been to use the individual student as the sampling unit, since this would have made it possible to draw accurate conclusions about American high school students in general on the basis of a smaller sample than with any other procedure. But from any realistic viewpoint this "ideal procedure" would have been patently impractical. Any procedure whereby a small fraction of the students in a class is taken out of the normal classroom routine for two days to take a battery of tests will almost certainly disrupt the

THE ORIGINAL DOCUMENT WAS OF POOR QUALITY. BEST POSSIBLE REPRODUCTION FROM COPY FURNISHED ASTIA.

learning schedule of the rest of the students in the class. Most school administrators feel that any large scale comprehensive testing program is most likely to be useful if results are available for all students in the school, or at least for all students in the grade, so that it will be possible to determine how any student stands in relation to his fellow-students--and also how the school as a whole stands in relation to other schools. Or to state the case another way, using the school as the sampling unit and testing all students within the selected schools has the following major advantages:

1. It is administratively convenient.
2. The results for the individual students are more useful, both to the student and to the school, when results are available for all students in the school.
3. It provides a better basis for research on the schools themselves--as opposed to research on the students. Certain kinds of questions, requiring comparisons of results for different kinds of schools, can be answered better when data are available for all students in the schools concerned than when only a small fraction of the students has been tested, even though a larger number of schools may be involved under the latter circumstances.

In brief, then, it was decided to use the school as the sampling unit, and to test all students in selected schools, not only because this seemed to be the practical procedure closest to the theoretically ideal but impractical procedure of using the student as the sampling unit, but also because it had theoretical advantages in its own right. (As things later turned out, it was agreed to depart from the chosen procedure in two places--New York City and Chicago--where the school administrators preferred to have more schools tested but only a fraction of the students in each school.

In New York City it was agreed to test one out of 12 students in every high school. In the case of Chicago's 38 academic and technical high schools, 20 of them were selected at random and one-tenth of the students in every grade in every selected school were tested.)

More specifically, in the case of the public schools, in situations where junior high schools existed, the sampling unit was to be the senior high school together with its associated junior high school(s). One purpose of treating junior high schools this way was to maximize the extent to which all four grades (9 through 12) would be included for the same school system. This special treatment of junior high schools was not deemed necessary in the case of private and parochial schools, where the junior high school is quite a rarity.

The reader will recall that stratified random sampling was the method that had been agreed upon, but that the decision as to how to stratify the sampling units could not be made until it had been decided what the sampling units would be. The decision that the sampling unit should be the school imposed the obvious requirement that the stratification variables must be variables applicable to the school. At the same time they were to be variables that there might be some reason to expect would be related to level of test scores of the students in the school. This relation might exist either because the chosen stratification variable was somehow related to the average ability level of the pupils or because it was related in some important way to the nature of the school program.

The primary stratification of the schools was on the

basis of whether they were public, parochial, or private.*

---

*As used in this report, these terms, "parochial school" and "private school", may require a little clarification. In the interests of brevity, the term "parochial school" is used to designate both parochial and diocesan schools under the auspices of the Roman Catholic Church. All other non-public schools, including those affiliated with churches other than the Roman Catholic, are designated "private school".

Public schools were then stratified on the basis of the following three variables:

1. School size (as represented by 12th Grade enrollment)
2. Location (primarily in terms of state)
3. Extent of school drop-outs

A word or two is in order on the way these three variables were handled, and the reasons for choosing them.

<u>School size</u>. School size was considered important because if a school is very small it cannot offer as varied a curriculum as a larger school, except at much greater cost per capita. This, incidentally, is one of the major theses of James Bryant Conant's widely quoted report on American high schools.*

---

*Conant, James B., <u>The American High School Today</u>. McGraw Hill, New York, 1959

Accordingly, the public high schools were divided into the following four strata in terms of Grade 12 enrollment (or Grade 11 in the case of those few school systems in the South which still have an eleven-year program):

1) Grade 12 enrollment: under 25
2) " " " : 25--99

3) Grade 12 enrollment: 100--399

4) " " " : 400 or more

One reason for splitting the groups in this particular way was so that two of the four groups would consist of schools large enough to meet the minimum size standard recommended by Conant* and the other two would fail to meet this standard.

Another advantage of this kind of split is that it would help insure a proper representation of urban and rural schools, since most of the extremely small schools (those in size category 1) would be in rural areas.

Location. Broad geographical region (e.g., New England, the Southwest, etc.) was an important consideration because of regional differences in the economy, the nature of the population, and the character of the schools. The actual state in which the school was located was also important because of the functions of most state departments of education in setting standards for the schools and the teachers, requirements for high school diplomas, etc. Another factor to be considered in connection with location was the distinction between very large cities and other types of local units, including smaller cities and villages. The geographical stratification procedure that was followed was one that had the advantage of taking into account these three important considerations--broad geographical region, specific state, and the distinction between very large cities and other local

---

*Conant, James B., The American High School Today. McGraw Hill, New York, 1959

units. What was done was to divide the country into the following 56 geographical units:

a. The five cities with populations in excess of 1,500,000 (New York, Chicago, Los Angeles, Philadelphia, Detroit).

b. The 50 states excluding the five cities mentioned above.

c. The District of Columbia.

These 56 geographical units were then grouped into nine broad geographical regions* ~~so that in the sampling process the "overflow" from one state would spill into a contiguous state rather than into a state that was merely alphabetically adjacent~~

Extent of school drop-outs. The proportion of students that drop out of high school before graduating is obviously a significant factor. A high drop-out rate is of course generally regarded as undesirable. A "retention index", representing the tendency for the students to remain in school until they graduate, was therefore computed for each high school, and used as a basis for stratification. The retention ratio was defined as the ratio of number of graduates in the '57-'58 school year to number of tenth-graders, in '58-'59.

The parochial schools and the private schools were stratified into the 56 geographical units, but unlike the public schools, they were not stratified on school size or retention ratio; this was because there were too few parochial and private schools to make it feasible to stratify them into so

---

*The nine regions are the same as those used by the U.S. Office of Education. The organization of the 56 geographical units of Project Talent into geographical regions is shown in Appendix 3A.

many subdivisions. The group of parochial schools and the group of private schools were sampled separately.

After the decision had been made that the school was to be the sampling unit, the next basic question was what "sampling ratio" should be used. The sampling ratio is the proportion of the "population" (or of any segment of the "population") that is selected for inclusion in the sample. Obviously, how high the ratio needs to be depends chiefly on how large the sample needs to be in order to yield dependable results in subsequent research based on the data.

In the case of Project Talent it was decided to aim at testing somewhere between 400,000 and 500,000 students. There were several reasons why so large a number seemed desirable. The reader will recall that the decision to use the school as the basic sampling unit necessitated a larger sample than would have been required if the individual student had been the sampling unit. Furthermore, it was realized that the initial sample would have to be extremely large because it would eventually be broken into so many different kinds of subdivisions that some of these smaller groups would be very small indeed. Considering only the initial breakdown, it was going to be necessary to obtain information separately for each grade, and within grades separately for each sex. But then information would also be needed for different sizes of schools, schools in different parts of the country, different types of curricula, and a vast number of other kinds of categories into which the students or schools could be classified immediately. It was not expected that the sample would include enough schools from each state to give usable information about individual states. Thus it

was explicitly planned <u>not</u> to get state norms. However, there did have to be enough students and schools included from each of the first eight* of the nine geographical regions (shown in Appendix 3A) to provide usable information about these larger areas. For instance regional norms were to be obtained.

In addition to the categories already mentioned in terms of which the students could be classified, there were other bases in terms of which it would be important to classify the students and analyze the data after the follow-up information had begun to come in. For instance the results would have to be analyzable separately for those students who go to college and those who do not; for those students who graduate from college and those who do not; for students who train for specific professions; for individuals who enter and attain at least moderate success in various occupations; and for individuals who achieve outstanding success in various ways; also, at the other end of the scale, for those students who become juvenile delinquents or adult criminals; for those who are chronically unemployed; for those who show other evidences of severe maladjustment.

In view of the fact that some of the occupations that are extremely important from the standpoint of the welfare of the nation (e.g., nuclear physicist) would ultimately attract only a very small percentage of the high school students tested, a very large number of cases would be necessary

---

*The ninth area, "non-contiguous states", was not covered by this requirement since Alaska and Hawaii were not states at the time the initial plans were set up.

in order to obtain enough cases in those important but relatively infrequent occupations that worthwhile information would be yielded about them. And even in the case of more usual professions, such as dentist, where the total number of persons would undoubtedly be larger, they might have to be divided into many sub-groups, homogeneous in terms of type of community, family background, high school curriculum, size of high school, sex of the student, the grade in which he was when the Project Talent tests were administered, and other factors. Since divisions of this sort produce such a very large number of sub-groups, the sub-groups themselves might be too small to provide useful information unless the initial size of the Project Talent group was very large indeed. This was perhaps the major consideration that led to the decision to aim at testing close to half a million students.

A second problem was the number of schools that would be needed. A large number of schools would be needed not only to insure an adequately representative group of students, but also because it was anticipated that in many of the special studies to be carried out with the Project Talent data, the schools, rather than individual students, would be the focus of concern. This made it important not only to have an adequate number of schools represented in the sample, but also to have adequate representation of each of the various <u>categories</u> of schools: urban schools, suburban schools, rural schools; large schools, small schools; private schools, public schools; schools in prosperous neighborhoods, schools in depressed areas; schools with conventional curricula, schools with experimental curricula; schools in the West,

schools in the East, schools in the South; academic high schools, commercial high schools, industrial high schools, comprehensive high schools. It was decided that the necessary representation of different kinds of schools could not begin to be achieved unless there were about 800 senior high schools in the sample, at the very least. Fortunately this number, 800 senior high schools, was compatible with the previously made independent decision that it would be desirable to test close to half a million students. Since information available from the Office of Education records suggested that the size of the average high school in the United States was somewhere in the vicinity of 500 students, it appeared that the 800 schools to be selected, together with associated junior high schools, would probably yield the necessary 400,000 to 500,000 students. This was a happy coincidence, but it did not solve all the problems.

One of the biggest problems it did not solve was the question of how to get an adequate representation of the larger public high schools. There are a great many more very small high schools in the United States than very large ones. If a uniform sampling ratio were used for all kinds of high schools, and if this sampling ratio were set to give the desired total number of high schools or the desired total number of students, the sample would be very heavily loaded with extremely small schools and would have far too few of the very large schools to make it possible to determine effectively the relationship between size of school and test results, or between size of school and subsequent activities of the student, or, for that matter, between size of school and anything else that might be of concern. The solu-

tion to this problem, the problem of having the public school sample contain too few very large schools and too many very small schools to be efficient for some types of analysis, lay in the use of differential sampling ratios. It was decided, therefore, to use a basic sampling ratio of one in 20 for medium-sized schools, a much larger sampling ratio, one in 13, for the largest schools, and a much smaller sampling ratio, one in 50, for the smaller schools. This was very easy to accomplish, since size of school (defined in terms of 12th Grade enrollment) was one of the variables which had been chosen as a basis of stratification. The sampling ratios used in most cases for the public senior high schools are summarized below:

School size stratum	12th Grade enrollment	Sampling ratio
1	under 25	1 : ~~13~~ 50
2 and 3	25 - 399	1 : 20
4	400 or more	1 : ~~50~~ 13

It was anticipated that this would provide about the right number of students with due allowance for expected absences. Exceptions to these specified sampling ratios in the case of New York City and Chicago have already been noted. The few other exceptions that occurred are described in Appendix 3D.

Naturally it was planned that the schools would be weighted appropriately in the analysis of the data, to adjust for the differential sampling ratios. To state the situation in its simplest terms, when a sampling ratio of one in 50 is used each school selected represents 50 schools; likewise each school selected on the basis of a one in 13 sampling ratio represents 13 schools. Therefore, in the data analysis these schools should be given weights of 50 and 13

respectively. Only by applying appropriate weights to the data can results be obtained which permit sound inferences about the total population of high schools or high school students. This matter of assigning school weights is discussed in a little more detail in Chapter VI.

For the parochial and private schools, on which no stratification except geographical had been done, a sampling ratio of one in 20 was used uniformly.

One of the complications in obtaining a representative sample of 9th Grade public school students was the fact that where junior high schools are an integral part of the school system, 9th Grade students are in a different school from the 10th, 11th, and 12th Grade students. This would create no difficulty if there were always a complete correspondence between junior high schools and senior high schools--in other words, if all graduates of a particular junior high school (or group of junior high schools) were expected to go to a particular senior high school and if all 10th Graders in that high school were expected to come from the junior high school (or group of junior high schools) in question. This convenient state of affairs exists in many communities, but in enough others to create technical problems in the sampling, the situation is more complicated; for instance some graduates of a junior high school may go to one senior high school and some to another. One solution, of course, would have been to treat the junior high schools just like the senior high schools, selecting a sample of them randomly within each stratum. But this was undesirable because it would have greatly reduced the number of school systems on which data were available for all four of the grades involved (Grades

9, 10, 11, 12). This could have seriously limited the usefulness of the data in special studies where the school, rather than the student, was the focus of concern; also studies where the pattern of differences between successive grades was of primary interest. It was decided, therefore, to attempt to select the junior high schools in such a way as to maximize their correspondence with the senior high schools already selected, while still maintaining due regard for the tenets of random sampling. Where junior high schools or groups of junior high schools were clearly and unambiguously associated with a specific senior high school that had been selected, those junior high schools were put in the sample. In communities where the situation was not so clear-cut, there was a problem of deciding what junior high school or group of junior high schools would include the maximum number of students who would go to the selected senior high school and the minimum number who would go to any other senior high school. At the sample selection stage this problem did not seem susceptible of mathematical treatment on any practical basis. It was therefore a matter of judgment. The necessary decisions were delegated to the appropriate Regional Coordinators--the persons connected with the project who were in the best position to obtain the needed information directly from the schools. Though the sample selection procedure for junior high schools was necessarily an approximative one, to a certain extent resultant inaccuracies could be corrected. This could be accomplished partly by means of a supplementary sampling phase in which a small number of junior high schools were chosen randomly from an appropriate group, and partly by means of appropriate adjustment

of the weights to be assigned to each of the junior high schools in the sample. The details of the supplementary sampling of junior high schools are presented in Appendix 3C. The problem of assigning appropriate weights to the schools is discussed in some detail in Chapter VI.

In connection with the way the schools were selected, there is one possible source of misunderstanding that requires clarification. It is extremely important to recognize that even though a school which is selected turns out to be located in some particular city or to have certain characteristics, the individual school is not necessarily representative of all the schools located in that city or having that characteristic; <u>nor is it expected to be</u>. Thus if one of the 16 schools in a certain city happens to be drawn it may not be a typical one. It might be the worst school in the city--or the best. It might be the largest, or it might be the smallest. But when combined with other schools selected in the same way in all parts of the country, it should contribute towards making the nation-wide sample a representative one. Thus its sole function in the sample is to contribute to the composite, along with many other schools, so that a clear picture can be obtained of the nature of American high schools and American high school students in general. If the total sample of schools is to give an accurate picture of the country as a whole, schools that are not typical of those in their locality have to have the same chance of being represented as do the more typical schools.

As has been mentioned in Chapter II, one important part of Project Talent was to collect information on one entire

age group. (Fifteen-year-olds happened to be the age group selected.) This meant the data had to be obtained not only for the 15-year-olds in high school but also for the ones not in high school. This would include, of course, 15-year-olds still in elementary school (Grade 8 or below), or in junior high school below the Grade 9 level, or already in college, or not in any school at all. The latter category would include those who had already graduated but had not entered college (such cases would be quite rare) and those who had dropped out of school without graduating. The non-high-school 15-year-olds included in the study were to be the residents of the areas served by one-tenth of the public senior high schools selected in the sample. For this reason, among others, the schools selected in the sampling phase were divided into ten sub-samples which in terms of the stratification variables would be as close to equivalent as could reasonably be achieved.* One of these ten sub-samples (designated "Sub-sample 0") was then selected to define the non-high-school 15-year-olds to be included in the study. Those included were all 15-year-olds not registered in <u>any</u> high school (public, parochial, or private) in Grade 9, 10, 11,

---

*Each senior high school was assigned to one of the ten sub-samples on the basis of an inspection of its characteristics in terms of the stratification variables. With some minor exceptions where missing data left no other alternative, the assignment was not random, since the purpose was to get ten sub-groups that would be as similar as possible. Each junior high school was assigned to the same sub-group as the senior high school with which it was associated--or in cases (such as New York City) where there was no direct association between junior high schools and specific senior high schools, the junior high school was assigned on whatever other basis seemed most reasonable.

or 12, who lived in the school district served by any of the public senior high schools in "Sub-sample 0". These non-high-school 15-year-olds are not intended to be considered part of the regular sample. However, when appropriately weighted they supplement the 15-year-olds in the regular sample (consisting of high school students) to provide a picture of the entire age group--the total group of 15-year-old Americans.

The Mechanics of Selecting the Schools

After the basic decisions on how to select the sample were made, it remained to actually carry out the selection procedure. This meant getting as comprehensive a list of the high schools in the United States as could be obtained, stratifying them on the agreed-upon basis, and then doing the random selection.

The comprehensive list of schools was put together from several different sources. The bulk of the information needed was provided by the United States Office of Education, which made available a record of the school names, locations, and other relevant information needed for stratification, in the form of a file of IBM cards, for most of the public senior high schools in the United States. This was supplemented by three lists of additional schools which the Office of Education later provided, and by a mailing list of schools that the Bureau of Internal Revenue provided, which happened to contain a few schools not on the Office of Education list. In addition, the Office of Education provided the names and addresses of the parochial and private high schools. The plan of operation was to arrange the schools in "sub-sets" of 13, 20, or 50, whatever number corresponded to the samp-

ling ratio for the type school in question. The schools within each sub-set were to be as similar as possible with relation to the appropriate stratification variables. One school was then to be selected at random from each sub-set. In the case of the vast majority of the public schools, those for which the Office of Education had provided IBM cards, it was possible to have the random selection actually carried out by an electronic computer, the "IBM-650". The machine was instructed, or to use the jargon of computer specialists, the machine was "programmed", to generate a series of random numbers, which designated which school in each sub-set of 13 or 20 or 50 was to be chosen. The IBM cards were fed into the machine, lights lit, wheels spun, circuits opened, circuits closed, vacuum tubes operated, the cards were "read", and computations were carried out at a fantastic rate of speed. The outcome of all this frenzied activity on the part of the "IBM-650" was a document printed by the machine, containing machine-generated instructions to the Project staff! This document was in the form of a list, specifying what schools were to be invited to participate in the project.

For the private schools, the parochial schools, and those public schools for which IBM cards were not available, the equivalent process had to be carried out by hand, with the aid of a statistician's tool known as a "table of random numbers". The Internal Revenue list and each of the supplementary lists provided by the Office of Education were treated separately in the sampling process. Since the Internal Revenue list provided no information about the school except address, this group could not be stratified, and

therefore a simple random sample of 2% was selected from it. There were a few other very minor departures from the agreed-upon sampling ratios. However, in all cases it was later possible to make appropriate adjustments in the school weights to correct fully for these discrepancies. For those who may be interested, the details are discussed in Appendix 3D, which also contains a somewhat more detailed description of the actual mechanics of the sampling procedure.

Response to Invitations to Participate

The preceding sections of this chapter describe at some length the many problems that had to be met and the many decisions to be made before the sample could actually be selected; there is also some hint that considerable time and effort went into the mechanics of actually selecting the schools. All this time and all this effort would be utterly in vain if any large proportion of the chosen schools were unwilling to cooperate. If this occurs, nothing can be done to salvage the results satisfactorily. Choosing other schools to invite as substitutes is no solution, for about the same reason that depending on schools to volunteer spontaneously is not a desirable method of sample selection. Just as schools that volunteer are probably not representative of those that do not, schools that decline to participate are likely to differ somehow from schools that agree to participate--and one can never know how much the results have been changed as a result. Fortunately, in the case of Project Talent this turned out not to be a major problem, since the schools recognized the importance of the Project and were eager to cooperate. Almost 93 percent of the schools selected in the sampling process agreed to participate--

really an astonishingly high percentage, and one which is a tribute both to the zeal of the Project's Regional Coordinators and to the vision of the school administrators, who recognized the immediate and long-term values of the Project.

Table III-1 shows the number of public, parochial, and private senior high schools invited to participate in each geographical unit (state or large city) and the number that accepted the invitation. For the public high schools the school size category is also shown. While retention index is not shown in this table a check was made on the basis of the public schools for which the Office of Education had provided IBM cards. This comprised almost all of the public schools. Among the schools selected from this group 44 had declined to participate. No evidence was found that these schools were biased to any substantial degree in regard to retention index.*

Table 1 makes it clear that the senior high schools gave remarkably good cooperation. In the junior high schools the picture is every bit as good, since there was not even one instance of a junior high school declining to participate if it had been determined by the regional coordinator to be an associated school for one of the senior high schools that had agreed to participate.**

Since the proportion of schools that declined to participate was so small, the problem of how to handle the situation was not really a serious one. It was decided that the best procedure would be to invite no additional schools as replacements, but instead to give a higher weight to the group of schools most similar to the missing school, in order to adjust the weighted number of cases so that it would

---

*The retention index percentile of the school within the group of schools in the same state and in the same school size category was determined. The median of these 44 percentiles was 45. Two thirds of the 44 percentiles fell between 14 and 89.

**There were, however, three refusals among the junior high schools in the supplementary junior high school sampling mentioned earlier in this chapter and discussed in Appendix 3C. This is hardly surprising, however, since the circumstances were such that these schools were notified very late in the school year.

Table III-1. Distribution of Sampling Units (Senior High Schools) that are Participating or Declined to Participate

Geographical Unit Code*	PUBLIC SENIOR HIGH SCHOOLS 1 - 24 Seniors			25 - 99 Seniors			100 - 399 Seniors			400 or more Seniors			TOTAL			PAROCHIAL HIGH SCHOOLS			PRIVATE HIGH SCHOOLS			ALL GROUPS COMBINED			PERCENT*** OF SCHOOLS	GEOGRAPHICAL UNIT CODE
Participating ⟶	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total		
1. New England....	4	-	4	16	-	16	12	-	12	1	-	1	33	-	33	13	-	13	9	-	9	55	-	55	100	1
111	2	-	2	3	-	3	1	-	1	-	-	-	6	-	6	1	-	1	2	-	2	9	-	9	100	111
112	-	-	-	2	-	2	1	-	1	-	-	-	3	-	3	1	-	1	1	-	1	5	-	5	100	112
113	1	-	1	2	-	2	-	-	-	-	-	-	3	-	3	1	-	1	1	-	1	5	-	5	100	113
114	1	-	1	7	-	7	5	-	5	1	-	1	14	-	14	8	-	8	3	-	3	25	-	25	100	114
115	-	-	-	-	-	-	2	-	2	-	-	-	2	-	2	-	-	-	-	-	-	2	-	2	100	115
116	-	-	-	2	-	2	3	-	3	-	-	-	5	-	5	2	-	2	2	-	2	9	-	9	100	116
2. Mid-east.......	3	-	3	45	1	46	66	4	70	54	3	57	168	8	176	26	5	31	10	3	13	204	16	220	92.7	2
2X1	-	-	-	3	-	3	32	-	32	50	-	50	85	-	85	3	1	4	2	-	2	90	1	91	98.9	2X1
221	3	-	3	15	1	16	9	2	11	-	1	1	27	4	31	9	-	9	3	-	3	39	4	43	90.7	221
222	-	-	-	2	-	2	5	2	7	-	2	2	7	4	11	4	-	4	2	-	2	13	4	17	76.5	222
2X3	-	-	-	-	-	-	1	-	1	1	-	1	2	-	2	-	2	2	-	1	1	2	3	5	40.0	2X3
223	-	-	-	19	-	19	17	-	17	2	-	2	38	-	38	6	2	8	2	1	3	46	3	49	93.9	223
254	-	-	-	2	-	2	-	-	-	-	-	-	2	-	2	-	-	-	-	-	-	2	-	2	100	254
255	-	-	-	4	-	4	2	-	2	1	-	1	7	-	7	4	-	4	1	-	1	12	-	12	100	255
250	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1	-	1	1	0.0	250
3. Great Lakes....	15	-	15	87	-	87	39	2	41	19	3	22	160	5	165	25	2	27	7	1	8	192	8	200	96.0	3
331	5	-	5	25	-	25	9	1	10	2	-	2	41	1	42	6	-	6	1	-	1	48	1	49	98.0	331
332	5	-	5	15	-	15	5	-	5	1	-	1	26	-	26	2	-	2	-	-	-	28	-	28	100	332
3X3	-	-	-	1	-	1	9	-	9	10	-	10	20	-	20	1	2	3	1	-	1	22	2	24	91.7	3X3
333	3	-	3	19	-	19	7	-	7	2	1	3	31	1	32	4	-	4	2	-	2	37	1	38	97.4	333
3X4	-	-	-	-	-	-	-	-	-	1	-	1	1	-	1	3	-	3	-	-	-	4	-	4	100	3X4
334	1	-	1	14	-	14	5	1	6	2	2	4	22	3	25	6	-	6	2	1	3	30	4	34	88.2	334
335	1	-	1	13	-	13	4	-	4	1	-	1	19	-	19	3	-	3	1	-	1	23	-	23	100	335
4. Plains.........	34	2	36	63	2	65	12	2	14	3	-	3	112	6	118	16	-	16	5	-	5	133	6	139	95.7	4
441	3	-	3	13	-	13	4	-	4	-	-	-	20	-	20	3	-	3	2	-	2	25	-	25	100	441
442	7	-	7	13	-	13	1	1	2	-	-	-	21	1	22	5	-	5	-	-	-	26	1	27	96.3	442
443	4	-	4	16	1	17	4	-	4	2	-	2	26	1	27	3	-	3	1	-	1	30	1	31	96.8	443
444	5	-	5	2	1	3	-	-	-	-	-	-	7	1	8	-	-	-	-	-	-	7	1	8	87.5	444
445	4	-	4	5	-	5	-	1	1	-	-	-	9	1	10	1	-	1	1	-	1	11	1	12	91.7	445
446	6	-	6	7	-	7	1	-	1	1	-	1	15	-	15	2	-	2	-	-	-	17	-	17	100	446
447	5	2	7	7	-	7	2	-	2	-	-	-	14	2	16	2	-	2	1	-	1	17	2	19	89.5	447
5. Southeast......	33	4	37	142	8	150	37	2	39	4	-	4	216	14	230	13	3	16	8	4	12	237	21	258	91.9	5
551	1	-	1	9	2	11	4	-	4	-	-	-	14	2	16	1	-	1	1	1	2	16	3	19	84.2	551
552	1	1**	2	6	-	6	3	-	3	-	-	-	10	1	11	1	-	1	-	-	-	11	1	12	91.7	552
553	4	-	4	26	-	26	3	-	3	-	-	-	33	-	33	1	-	1	2	-	2	36	-	36	100	553
554	2	-	2	10	1	11	3	-	3	-	-	-	15	1	16	-	-	-	-	-	-	15	1	16	93.8	554
555	3	-	3	12	2	14	4	1	5	-	-	-	19	3	22	1	-	1	-	-	-	20	3	23	87.0	555
556	2	-	2	6	-	6	4	-	4	2	-	2	14	-	14	1	-	1	-	2	2	15	2	17	88.2	556
561	2	-	2	14	-	14	2	-	2	1	-	1	19	-	19	3	1	4	2	-	2	24	1	25	96.0	561
562	2	-	2	12	1	13	5	-	5	-	-	-	19	1	20	-	-	-	1	-	1	20	1	21	95.2	562
563	4	1	5	15	1	16	2	1	3	1	-	1	22	3	25	-	-	-	1	-	1	23	3	26	88.5	563
564	3	1	4	10	1	11	1	-	1	-	-	-	14	2	16	2	-	2	-	-	-	16	2	18	88.9	564
571	5	1	6	10	-	10	2	-	2	-	-	-	17	1	18	-	1	1	-	-	-	17	2	19	89.5	571
572	4	-	4	12	-	12	4	-	4	-	-	-	20	-	20	3	1	4	1	1	2	24	2	26	92.3	572
6. Southwest......	24	1	25	35	5	40	10	1	11	4	-	4	73	7	80	7	-	7	4	-	4	84	7	91	92.3	6
671	7	-	7	11	-	11	2	1	3	1	-	1	21	1	22	1	-	1	1	-	1	23	1	24	95.8	671
672	15	1**	16	21	5	26	6	-	6	3	-	3	45	6	51	4	-	4	2	-	2	51	6	57	89.5	672
683	2	-	2	2	-	2	1	-	1	-	-	-	5	-	5	1	-	1	1	-	1	7	-	7	100	683
684	-	-	-	1	-	1	1	-	1	-	-	-	2	-	2	1	-	1	-	-	-	3	-	3	100	684
7. Rocky Mountains	8	-	8	11	2	13	3	2	5	2	-	2	24	3	27	2	-	2	2	-	2	28	3	31	90.3	7
781	3	-	3	2	2	4	1	-	1	1	-	1	7	2	9	-	-	-	-	-	-	7	2	9	77.8	781
782	-	-	-	2	-	2	-	1	1	-	-	-	2	1	3	1	-	1	1	-	1	4	1	5	80.0	782
783	1	-	1	1	-	1	-	-	-	-	-	-	2	-	2	-	-	-	-	-	-	2	-	2	100	783
784	4	-	4	4	-	4	-	1	1	1	-	1	9	-	9	1	-	1	-	-	-	10	-	10	100	784
785	-	-	-	2	-	2	2	-	2	-	-	-	4	-	4	-	-	-	1	-	1	5	-	5	100	785
8. Far West......	3	1	4	14	2	16	13	5	18	4	6	10	34	14	48	12	1	13	5	-	5	51	15	66	77.3	8
881	-	-	-	-	-	-	1	-	1	-	-	-	1	-	1	1	-	1	-	-	-	2	-	2	100	881
892	1	1	2	5	-	5	3	-	3	-	-	-	9	1	10	2	-	2	2	-	2	13	1	14	92.9	892
893	2	-	2	4	-	4	2	-	2	-	1	1	8	1	9	1	-	1	1	-	1	10	1	11	90.9	893
8X4	-	-	-	-	-	-	1	-	1	1	2	3	2	2	4	1	1	2	-	-	-	3	3	6	50.0	8X4
894	-	-	-	5	2	7	6	5	11	3	3	6	14	10	24	7	-	7	2	-	2	23	10	33	69.7	894
9. Non-contiguous	-	-	-	-	-	-	2	-	2	-	-	-	2	-	2	-	-	-	1	-	1	3	-	3	100	9
901	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	—	901
902	-	-	-	-	-	-	2	-	2	-	-	-	2	-	2	-	-	-	1	-	1	3	-	3	100	902
TOTALS	124	8	132	413	20	433	194	18	212	91	12	103	822	57	879	114	11	125	51	8	59	987	76	1063	92.9	

---

*Refer to Appendix 3A for the names of states and special geographic units.

**Internal Revenue list, sampling ratio 1/50.

***Percent of selected sampling units that agreed to participate.

correspond with the number of cases in the "population".

The other aspect of the sampling, the identification and testing of 15-year-olds not in high school who were living in designated school districts, also met with good cooperation.

Although exact counts are not available yet, and will not be until a later stage in the data processing is reached, it would appear that all but a couple of hundred 15-year-olds in the designated school districts who are not in high school were tested.

Special Groups

In addition to the students in Grades 9 through 12 in the junior and senior high schools that constituted the regular sample, other boys and girls were also tested. Included in this group, of course, were the 15-year-olds who were not in high school; they have already been discussed. But in addition to them there were other cases, tested either because they were needed for inclusion in one of the special studies that are contemplated, or in a few cases chiefly as a matter of convenience. Since the students in such schools cannot be considered part of the regular sample, the data from them will not be included routinely in all studies; such schools will be included only in those special studies in which they are appropriate.

The reasons for testing these various groups of boys and girls who are neither in the regular sample of Grade 9 through 12 students nor in the sample of 15-year-olds are many and varied. However, a relatively small number of categories account for most of these additional cases.

For instance, there were a few places where one school

in a community was selected in the random sampling and the school authorities wished to test the other local schools as well. The desired arrangements were made, but since the additional schools were not drawn in the random sampling process, they are not considered part of the regular sample; the data from them will therefore be included only in those special studies in which it is appropriate to do so.

Then there are the eighth-graders. Many of the 15-year-olds not in high school turned out to be in elementary school in Grade 8. Rather than disrupt a class which had many 15-year-olds by giving the Talent battery to only a fraction of the class, a substantial number of the elementary schools decided to test their entire eighth grade--all the pupils, regardless of their age.

Then there is Knoxville, Tennessee, and the surrounding county, Knox County. Two schools were drawn as part of the regular sample in this area, but as a result of the enthusiasm of the school authorities there it was possible to arrange to test every student in every school--public, parochial, and private in the entire Knoxville and Knox County area, not only Grades 9 through 12, but also Grade 8. This very comprehensive testing in a concentrated area, over a five-grade range, will make possible many special studies that could not otherwise be carried out.

And last but by no means least is a small group of selected schools which are among those that have been trying out experimental curricula in mathematics. Two experimental curricula are involved here--one developed by Dr. E. G. Begle of Yale University and one by Dr. Max Beberman of the University of Illinois. It is planned to undertake some

special studies of the Talent data, to compare the performance of students in these experimental curricula with that of matched groups who take the conventional kind of high school mathematics courses.

Appendix 3A. List of the 56 basic geographical units*, classified into regions **

1. New England (6 geog. units)

111	Maine
112	New Hampshire
113	Vermont
114	Massachusetts
115	Rhode Island
116	Connecticut

2. Mid-east (8 geog. units)

2X1	New York City***
221	New York (except N.Y.C.)
222	New Jersey
2X3	Philadelphia***
223	Pennsylvania (except Phila.)
254	Delaware
255	Maryland
250	District of Columbia

3. Great Lakes (7 geog. units)

331	Ohio
332	Indiana
3X3	Chicago***
333	Illinois (except Chicago)
3X4	Detroit***
334	Michigan (except Detroit)
335	Wisconsin

4. Plains (7 geog. units)

441	Minnesota
442	Iowa
443	Missouri
444	North Dakota
445	South Dakota
446	Nebraska
447	Kansas

5. Southeast (12 geog. units)

551	Virginia
552	West Virginia
553	North Carolina
554	South Carolina
555	Georgia
556	Florida
561	Kentucky
562	Tennessee
563	Alabama
564	Mississippi
571	Arkansas
572	Louisiana

6. Southwest (4 geog. units)

671	Oklahoma
672	Texas
683	New Mexico
684	Arizona

7. Rocky Mountains (5 geog. units)

781	Montana
782	Idaho
783	Wyoming
784	Colorado
785	Utah

8. Far West (5 geog. units)

881	Nevada
892	Washington
893	Oregon
8X4	Los Angeles***
894	California (except Los Angeles)

9. Non-contiguous (2 geog. units)

901	Alaska
902	Hawaii

---

* The 3-digit number preceding the name of the geographical unit is the code number assigned to it for use in machine processing.

** The nine regions shown, and indicated by the first digit of the geographical unit code, are the regions used by the U.S. Office of Education. The second digit of the geographical unit code, except where it is X, indicating a city with population over 1,500,000, indicates the geographical regions used by the Bureau of the Census.

*** Cities over 1,500,000.

APPENDIX 3B

Miscellaneous Notes on the Sampling

After the sample was selected and the testing completed, certain minor inaccuracies in the sampling procedure, due chiefly to misinformation at the time the sample was selected, came to the attention of the project staff. While it is believed that these inaccuracies are, in general, too trivial to have any significant effect on the results for most purposes, they are listed here to permit the reader to judge for himself.

1. The Chicago sample consisted of 20 schools drawn randomly from a list of 38 schools which were believed to constitute the total group of Chicago secondary schools. It was later discovered that these 38 schools were all academic and technical high schools, and that in addition, there were nine vocational high schools which had not appeared on the list, and which therefore had no opportunity to be represented in the sample.

2. It will be recalled that in the case of the public schools, most of the actual sample selection was performed by an electronic computer. An anomaly in the machine set-up had the result that whatever school happened to correspond to the very last card in the sub-group (of 13 or 20 or 50 cards) had only half as much chance of being selected as each of the other schools in the sub-group.

CHAPTER IV

HOW THE TESTS WERE CONSTRUCTED

The decision on the contents of the battery of tests was a major one because on it depended how almost a million pupil-schooldays were to be spent (two days for each of almost a half-million pupils).

When a new test battery is to be put together, there are two reasonable ways of doing it. One is to select a number of tests from among the many fine ones already available, including both commercially published and non-commercial tests. The other is to develop brand-new tests. Very serious consideration was given to the former possibility, but it was rejected on numerous grounds. The chief reason was that tests not specifically constructed with the purposes of Project Talent in mind could not be optimally satisfactory for meeting those purposes.

To be optimally satisfactory a test battery has to be designed with the special purposes for which it is intended in mind. Each test must make its own unique contribution to those purposes. A test, however good it may be when used alone, serves no useful function in a battery if it accomplishes nothing in the way of measurement that is not duplicated in effect by some other test or combination of tests already in the battery.

The Problem of Test Length

Then there is the rather troublesome problem of test length. A ready-made test will seldom turn out to be the best length for the purpose it is to serve in the battery. It will probably come as no surprise to the reader that the test must be long enough to give a reasonably accurate measure, but it may surprise some readers to learn that the use of tests that are not short enough is also a hazard. If it is important to measure the pupils' ability to spell, and if it is all right to have 16 test questions devoted to this purpose why would it not be ten times as good to have 160 items on spelling? Surely high school students ought to know how to spell more than 160 words correctly. Surely, then, it would be reasonable to test spelling ability on as many as 160 words. Surely the results obtained from 160 test items would give a more accurate picture of the students' spelling ability than could be obtained from a mere 16 items. Yes, surely! But if 16 items take eight minutes of testing time, 160 items would eat up 80 precious minutes. And the 72 extra minutes devoted to the measurement of spelling ability would mean that the measurement of several other important kinds of aptitude or ability would have to be left out because there would not be time enough left for them. Thus it is just as important to have the tests short enough as to have them long enough. Tests that are to be used for survey purposes and in combination with other tests to provide composite scores can reasonably be much shorter than tests which are to be used for diagnosis and tests whose scores are to be used individually. For purposes of prediction a large number of fairly

short tests is generally preferable to a small number of longer tests. While individual scores on the latter may be more accurate (or to use the technical term, more "reliable") the resultant composites will probably be much less predictive. Thus one major advantage of the decision to develop a battery from scratch, specifically for Project Talent, was that each test could be made as long or as short as seemed desirable.

There were other important advantages, too, to be derived from having new tests, designed specially for Project Talent. One was that this insured that none of the tests would ever have been taken before by any of the students in the Project. A somewhat related advantage lay in the fact that because the tests were specifically and solely for Project Talent, they could be taken out of circulation. The tests would not be available for future use, except for studies related directly to Project Talent, or for calibration of other standardized tests so that the Project Talent results could be applied to those tests.

Once the decision had been made to develop a special battery of tests for Project Talent, it meant that a decision had also been made to undertake a long chain of intermediate steps. This chain would link the initial planning phase to the end-product, the final test battery. The general pattern of these steps is outlined below.

Table IV-1: Composition of Experimental Battery

Experimental Battery							Test Name in Final Battery
Test #	Test Name	Options per item	Mins. per form	No. of items Form A	Form B	Total	
1.	Vocabulary-Information Profile	5	170	355	354	696*	Information
2.	English: Active Vocabulary	5	8	15	15	30	****
3.	Effective Expression	2	9	10	10	20	***
4.	English Usage	2-5	14	24	24	48	***
5.	Sentence Structure	3	6	8	8	16	Punctuation.Section b. {Sentence Structure
6.	Punctuation	2-5	12	23	23	46	a. {Punctuatio Marks
7.	Capitalization	2	15	71	71	71**	***
8.	Spelling	5	14	18	18	36	***
9.	Reading Comprehension	5	26	51	51	102	***
10.	Following Directions	5	12	15	15	30	****
11.	Disguised Words	5	3	25	25	50	***
12.	Words in Sentences	5	12	20	20	40	Word Functions in Sentences
13.	Paired Associates	5	3	12	12	24	Memory for Words
14.	Sentence Completion	5	5	8	8	16	Memory for Sentences
15.	Arithmetic Computation	5	20	60	60	60**	***
16.	Arithmetic Reasoning	5	28	18	18	36	Mathematics I. Arithmetic Reasonin
17.	Mathematics A. (through Grade 9)	5	40	19	19	38	II. Introductory
18.	B. (Grades 10-12)	5		16	16	32	III. Advanced
19.	Verbal Reasoning	3	14	25	25	50	****
20.	Abstract Reasoning	5	14	18	18	36	***
21.	Mechanical Reasoning	2-5	12	24	24	48	***
22.	Spatial A. Folding	5	8	18	18	36	Visualization in Two Dimensions
23.	Spatial B. Rotation-Reflection	5	4	16	16	32	Visualization in Three Dimensions
24.	Scale Reading	5	6	24	24	48	****
25.	Name Comparison	2	6	100	100	100**	Clerical Checking
26.	Table Reading	5	3	70	70	70**	***
27.	Form Perception	5	3	36	36	72	Object Inspection
28.	Creativity	5	16	15	15	30	***
29.	Social Judgments	2	2	90	90	90**	Preferences
	Total			1204	1203	2003	

~~* 13 items (screening scale) are the same in both forms~~

** Same items in both forms

*** Same test name in final battery as in experimental battery

**** Not in final battery

A. Planning the experimental battery

   1. Determining the broad outlines of the battery--the general form it would take, the restrictions and limitations it would have to conform to, the functions it would have to serve.
   2. Review of past research and experience, to determine what kinds of tests to develop for the experimental battery. (It is desirable to include more kinds of tests in the experimental battery than it is expected will be used in the final battery.
   3. Preparing detailed specifications for each test.

B. Developing the tests of the experimental battery

   1. Writing the items and organizing the tests in accordance with the specifications.
   2. Putting the tests together into an experimental battery, and preparing the necessary accessory materials.

C. Tryout of the experimental battery

   1. Arranging for tryout in a suitable group of schools.
   2. Actual tryout.
   3. Reporting results to the schools.

D. Analysis of the tryout results

   1. Initial processing of the results.
   2. Scoring of the tests.
   3. Statistical analysis of the scores to decide which tests to eliminate from the battery.
   4. Item analysis; in other words, statistical analysis of the results for each test item.

E. Development of the final battery

1. Development of the final forms of the tests that are to be retained in the battery.
2. Organization of the tests into an integrated battery and development of the necessary accessory materials.

F. "Dress Rehearsal"

1. A dress rehearsal of the final battery to see that nothing has been overlooked.
2. Making whatever minor revisions (in directions and timing) seem desirable.

That is a standard pattern for developing a test battery, and it is the pattern that was followed in the case of Project Talent. Now let's go back to the beginning--planning the broad outlines of the battery.

Planning the Battery

What were the limitations on the kinds of tests to be tried out? First of all, there was a time limit; the schools in the try-out group could not be expected to provide more than two days of testing time. Secondly, the tests had to be of the objective type, with responses marked on an answer sheet that could be scored by a machine. The practical considerations that dictated the use of objective tests are of course obvious, especially in view of the magnitude of the test-scoring operation even under the best of circumstances. But even if there had not been this very practical reason for requiring objective tests, they still would probably have been the method of choice, because of their greater "reliability". Thirdly, most of the tests had to be of reasonably well established types, since this was to

represent a "state-of-the-art" battery. In other words the development of the battery was to be based chiefly on what had been learned over the years about the art of test construction and about what various kinds of tests measure and predict, rather than on mere speculation. Most of the tests thus had to measure abilities, aptitudes, and kinds of achievement of demonstrated importance and predictive value. While tests of a frankly experimental nature were not ruled out, they could certainly represent no more than a very small fraction of the final battery.

With these limitations in mind, decisions as to what tests would be constructed and tried out could be made on the basis of a review of past research concerning what sorts of measures are useful for what purposes*. Recommendations by the project's Test Panel were of course given heavy weight in making the decisions.

As has already been mentioned, in developing a battery it is not unusual to use a larger number of tests in the experimental form than it is expected will be used in the final battery. This is done with the expectation that some tests will be eliminated on the basis of the resultant data. Project Talent followed this standard practice, constructing about 25 percent more tests than it seemed likely the final battery would include. In view of the fact that tests would be discarded from the battery eventually, the chief concern was to get an adequately comprehensive battery for tryout purposes. In this initial stage it was more important to avoid errors of omissions than errors of commission.

---

* The historical background of each of the types of tests used in Project Talent is presented in Chapter VII.

It was decided to develop experimental forms of 29 tests, plus three inventories (an interest inventory, a student activities inventory, and a student information blank). Most of the rest of this chapter is concerned with the tests rather than with the three inventories, since construction of the inventories is discussed in detail in Chapter VIII.

The 29 experimental tests were designed to measure aptitudes, achievement in various areas, and acquisition of various kinds of information. The names of the tests are listed in Table IV-1.*

Developing the Tests of the Experimental Battery

Before the tests themselves could be written, detailed specifications had to be prepared for them, stating not only the number and kind of items but also the reasons for using items of that type, suggestions for constructing them, characteristics the items should have, characteristics they should not have, and other details. Underlying all this specification-writing was the philosophy that on the quality of the specifications depends the quality of the final instruments. This philosophy has sometimes been expressed in the statement that most of the time allocated to developing a test should be devoted to preparation of the "rationale" for the test--in other words, determining what to measure and how to do it. If this is done thoroughly and if the items written conform to the specifications prepared, the resultant test is far more likely to be a good one than than if the test author dived headfirst into the item-writing

---

* For some of the tests the name used in the final battery differs from that in the experimental battery. In such cases Table I shows both test names.

phase without any preliminary research or planning.

As in the determination of the number of experimental tests, in the determination of the length of the individual tests it is customary to be somewhat more generous than will be feasible in the final battery. Project Talent followed this customary procedure. Writing more items than one expects to use in the final form permits elimination of some items of the experimental form on the basis of the tryout data.

With very few exceptions the specifications called for tests which were not strictly parallel to any other published tests. While some of the item types were somewhat similar to rather widely used ones, the tests involved usually differed in length, distribution of content, and other details.*

The exceptions were two tests already published which involved novel elements that seemed sufficiently promising to warrant development (with the authors' permission, of course) of essentially parallel forms for the Talent battery. One of these two tests was the Psi Lambda Test ( ) developed by Dr. John B. Carroll and Stanley M. Sapon. This was a test to predict aptitude for learning foreign languages. Three of the tests in the Talent battery, Disguised Words, Word Functions in Sentences, and Memory for Words, are based closely on Subtests 3, 4, and 5, respectively, of the Psi Lambda Test.**

---

* The specifications for each of the Project Talent tests and the details as to the extent to which these tests differ from previously used somewhat similar tests are discussed in Chapter VII.

** The Psi Lambda Test has since been published by The Psychological Corporation under the title "Modern Language Aptitude Test" ( ).

The other test on which a test of the Talent Battery was modeled directly was the Ingenuity Test of the Flanagan Aptitude Classification Test (7) battery. The Creativity Test of the Talent Battery is essentially a parallel form of this test.

Although the final forms of the individual tests are discussed in considerable detail in Chapter VII, a brief description of each experimental test is presented below.

1. Information Test. The experimental form yielded a total score and 24 subscores, indicating information in 24 separate areas.
2. Active Vocabulary. A test of ability to find (rather than merely to choose from several options) the precise word fitting a given definition.
3. Effective Expression.
4. English Usage.
5. Sentence Structure.
6. Punctuation.
7. Capitalization.
8. Spelling.

These six subtests comprise the English Test

9. Reading Comprehension. A test of ability to comprehend written materials dealing with various types of subject matter.
10. Following Directions. A test of ability to comprehend and follow simple printed directions.
11. Disguised Words. A test of ability to decipher words whose spelling is distorted. (This is intended to measure one kind of aptitude involved in learning a foreign language.)
12. Word Functions in Sentences. A test of sensitivity to grammatical functions of words.
13. Memory for Words. A test of ability to memorize foreign words corresponding to common English words.
14. Memory for Sentences. A test of one kind of memory--the ability to memorize simple sentences and recall a missing word.

15. Arithmetic Computation. A test of speed and accuracy of computation.

16. Mathematics I (Arithmetic Reasoning). A test of ability to do the kind of reasoning required to solve arithmetic problems.

17. Mathematics II (Introductory). Mathematics (other than the kinds measured by Tests 15 and 16 above) through the 9th grade level.

18. Mathematics III (Advanced). College-preparatory mathematics generally taught in Grades 10-12.

19. Verbal Reasoning. A test of ability to do the kind of reasoning involved in verbal syllogisms.

20. Abstract Reasoning. A non-verbal test of one kind of abstract reasoning ability--the ability to determine logical relationships among the elements of a pattern of diagrams.

21. Mechanical Reasoning. A test of ability to visualize the effects of everyday physical forces and principles (for instance, gravitation) and the operation of basic kinds of mechanisms (for instance, gears, pulleys, wheels, springs, levers).

22. Visualization in Three Dimensions. A test of ability to visualize how a two-dimensional figure would look after it had been folded to make a three-dimensional figure.

23. Visualization in Two Dimensions. A test of ability to visualize how diagrams would look after being turned around on a flat surface, in constrast to how they would look after being turned over.

24. Scale Reading. A test of ability to read simple dials and scales quickly and accurately.

25. Clerical Checking. A test of speed and accuracy of perception in a simple clerical task, involving comparison of names.

26. Table Reading. A test of speed and accuracy in a non-computational clerical task involving obtaining information from tables.

27. Object Inspection. A test of speed and accuracy in perception of form.

28. Creativity. A test of the ability to find ingenious solutions to a variety of practical problems.

29. Preferences. A test involving a decision-making task. (This instrument is still in the experimental stage.)

Because only two days of testing time would be available for any student in the tryout, and because it was desirable to try out more than two days' worth of tests and test items, two parallel forms of most of the tests were developed (designated Form A and Form B respectively). Since each student was to be given only one of these two forms, it meant that more items could be tried out in the available two days.

In the case of most of the tests, each of the two experimental forms was set up with two separate equivalent halves, to be administered with separate time limits. The reasons for this will be immediately apparent to specialists in the field of test construction. For other readers it will probably suffice at this point to say that the reason was that it was necessary to

be able to find out in the experimental tryout to what extent the tests were successful in providing meaningful scores (as opposed to scores which represented random variation, not real measurement). This is discussed in a little more detail later in this chapter, under the heading "Test Reliability".

The items developed for the battery were all brand new, and developed specially for the purpose by the Project staff, except in the case of three of the tests (Scale Reading, Visualization in Three Dimensions, and Object Inspection, for which the United States Civil Service Commission very kindly provided some previously unused materials.*

Experimental Tryout and Data Analysis

The experimental battery was tried out in 11 widely scattered high schools. In deciding what high schools should participate in this tryout, considerable importance was attached to the desirability of having a wide variety of types of schools represented. By design, the group of eleven schools finally selected for this tryout included schools in large cities, schools in small cities, schools in suburban areas, and schools in rural areas; large schools, medium-sized schools, and small schools; schools in the Northeast, schools in the South, and schools in the Midwest. All students in Grades 9-12 in these schools were tested. This was a total of nearly 6,000 students. They each filled out six IBM Answer Sheets.

---

* For the Object Inspection Test and the Visualization in three Dimensions Test the United States Civil Service Commission provided some of the items. For the Scale Reading Test the Commission provided diagrams which were used as a basis for the items. The Commission also contributed the services of a draftsman in preparing some of the diagrams for the Mechanical Reasoning Test.

The purposes of the tryout have already been implied but perhaps it would be well to state them explicitly at this point, so that the reader will understand what was necessary and why. The tryout had two main purposes. The first was to decide what tests were to be in the final battery and how long they should be. In other words, the basic question to be answered was: "How can the available time best be apportioned among the available kinds of tests?" The second purpose was to provide a basis for deciding in the case of those tests that were to be in the final battery, what test items should be included. The reader will recall that more items had been written than it was expected would be used in a final form of the test. Therefore many items could be eliminated on the basis of the item analysis. Some of the items that were to be retained might need minor revisions; the item analysis would be helpful here too. For both of the purposes mentioned above--the determination of the composition of the final battery in terms of the tests in it and determination of the composition of each of the tests in terms of its items--it would be necessary to know the test scores. They would be needed for a third purpose too--so that test results could be reported to the schools that participated in the tryout.

The procedure for the initial processing of the results was quite novel. Contrary to the customary procedure, scoring was not the first step. Furthermore when the score was eventually carried out, it was not done either by hand or by a scoring machine. Instead it was done by a general purpose electronic computer. Since it was going to be necessary to have information not only about each student's

score on each of the tests but also about his specific response to each test item, a procedure was used for getting all of a student's responses transferred directly from his answer sheets to a set of IBM cards, in which the specific responses were represented by punched holes. A special machine developed by Mr. Elmer Hankes, of Testscor, Inc., in Minneapolis, made it possible to transfer these responses directly from the answer sheets to the cards without any human intermediary (except to set up the machine, of course). Answer sheets and blank cards were fed into the machine; answer sheets and punched cards came out. There were 12 IBM cards for each student, or a total of almost 70,000 cards. These cards were then fed through a different machine, the IBM 704 Computer, which provided "item analysis" data--the information on how the students responded to the test items--in order to decide which items to retain and which items to eliminate. At the same time the computer actually scored the tests, obtaining a total of 53 scores for each student.

Some of the readers may be puzzled by the concept of item analysis. They may ask, "How can a bunch of numbers tell you whether your test items are any good?" The answer is that actually they can't. Item analysis is no substitute for judgment and care in writing the items; nothing is. But item analysis can be an invaluable supplement. Let's see how it works.

Consider the following test item (which we hasten to point out is not one that was ever really used) for a general information test.

A mixture of blue and yellow produces

- A. orange.
- B. red.
- C. green.
- D. white.
- E. brown.

This is a very bad test item because it is ambituous; it has two defensible answers instead of just one. The physicist knows that a combination of blue light and yellow light produces white; the artist knows that mixing blue pigment with yellow pigment produces green pigment. Obviously, then, either C or D is a defensible answer. Now let us suppose that the author of the item does not know much about physical science, and that he therefore thinks C is the only answer. When he prepares the key for the test, he makes C the key answer for that item. The students who know about neither the mixing of pigments nor the mixing of colored light rays will have no idea of the answer. They may guess any of the choices or they may omit the item. The students who know about pigment-mixing but not about light-mixing will pick C. The other group of students with partial knowledge, those who know about light-mixing but not about color-mixing, will pick D. But what about those students who are best informed? They are the ones who will really be puzzled by the item since they may recognize that it is ambiguous. These students may answer C, or they may answer D, but they are likely to omit the item entirely. Now suppose that one part of the item analysis is to see what kind of total scores on the test the students who pick each answer get. The results are quite likely to be

somewhat as follows:

The students who pick A, B, or E, will have lower total scores on the average than those who pick C or D; but those who omit the item altogether will be likely to have the highest scores of all. These would not be the sort of results one would expect for a good item, and therefore they should serve as a signal that something is amiss. A further clue as to the nature of the difficulty might be provided if, as seems likely, the average score of the students who chose D were about as high as the score of those who picked the supposed key answer, C. If further delving revealed exactly what the trouble is, the item could be revised and salvaged. For instance the difficulty could be obviated by changing the item as follows:

Mixing blue paint with yellow paint produces

A. orange paint.
B. red paint.
C. green paint.
D. white paint.
E. brown paint.

Then the item would have only one correct answer, C, (although option D in the new version might be more plausible, and therefore better, if it were changed to violet).

Now let us look at another very bad test item--or at least one which would be bad for most purposes. Suppose that the following item is included in a ninth-grade mathematics test:

The value of π is closest to

A. 3.141588

B. 3.141589

C. 3.141591

D. 3.141593

E. 3.141594

This is an absurd item. It would be utterly ridiculous to expect 9th grade students to have memorized the value of π to that many decimal places. As a matter of fact it would be ridiculous to expect it of a professional mathematician. However, if this item did get into a mathematics test the results for the item analysis would probably indicate that good and bad students alike were apparently guessing on the item. Probably about the same number of students would choose each option, and those who chose option D, the key answer, would not have any higher scores on the rest of the test than those who chose A, B, C, or E. This item would undoubtedly be thrown out of the test, since the item analysis data would make it immediately apparent that it was too hard. A test item to which no one in the group for which the test was intended knows the answer has no place in the tests.

Likewise item analysis results can point out items that are too easy.

Each response to each item in the Talent battery was analyzed separately for the lower grades of high school (grades 9 and 10), and the upper grades (11 and 12). The information obtained for each option of each item included the following:

1. The percentage of those students who reached the item that select the option.

2. The average score of the students who selected that option.
3. A value called the "point biserial correlation coefficient" which indicates the degree to which students who pick the response in question tend to have higher (or lower) scores than students who pick other responses for the same item.

In addition the following information was obtained for each item:

1. The percentage of students who reached the item.
2. The percentage of those students who reached the item who deliberately omitted it.
3. Their average score obtained on the test by the students who deliberately omitted the item.
4. The corresponding point biserial correlation coefficient.

The reason for obtaining the latter three values was so that in the item analysis process, deliberate omission of an item could be treated as what it really was--a form of response.

The percentage of students reaching each item was determined for an entirely different reason--to provide a basis for deciding on time limits for the tests, or alternatively for deciding on the number of items appropriate for a given time limit.

In doing an item analysis on a test battery as large as this, a monumental mass of data is the outcome. In this instance there were two forms of a 29-test battery,

including a total of 1,518 items in all, with a total of almost 10,000 options; this meant that item analysis data of the type indicated had to be obtained for a total of approximately 12,000 options.

The item analysis provides the basis for revising the experimental forms of the tests to produce final forms, if the test is to be included in the final battery. But the item analysis alone does not make a very good basis for the decision as to whether the test should be included. Additional data are needed for that decision. The main problem is one of efficient use of the invaluable testing time. This means that the final battery should consist of tests that will measure as many different things as possible and will overlap each other as little as possible. A second consideration, closely related to the first, is whether the tests are adequately reliable; in other words, whether two parallel forms of the test would give consistent results.

At this point still another electronic computer entered the picture--the IBM-709, which somewhat resembles the 704 but is bigger, faster, and if an anthropomorphic description were to be used, even "smarter".

To answer the question about test reliability, statistical values called "reliability coefficients" were computed. Reliability coefficients represent the correlation* between parallel forms of the test. This correlation coefficient may be computed directly if parallel forms of the test are available; otherwise it may be computed indirectly in any one of several possible ways. One of these

---

* Footnote on next page

ways, which involves determining the relationship between scores on two halves of the same test, the halves being separately timed, was the method used for most of the experimental forms of the Talent tests.

A test that has any reliability at all can be made more reliable by adding more items. Thus the reliability coefficient helps determine not only whether a test should be retained but how many items it should have.

Knowing that a test has satisfactory reliability is fine, but it is still not enough. To decide what tests to pick in order to minimize overlap and maximize coverage it is necessary to know about the interrelations among the tests. Therefore intercorrelations were computed among all 53 scores, and these values, together with the reliability coefficients, provided the basis for subsequent computation of indices which provided a measure of the extent of each test's unique contribution to the battery. The details of this subsequent computation are presented in Appendix 4A.

Development of the Final Battery.

On the basis of careful consideration of all available facts, it was decided to eliminate four tests that were in

---

* A correlation coefficient is a number representing the degree of relationship between two variables (scores on two different tests, for instance). Correlation coefficients can have any value between +1 and -1. A high positive correlation coefficient indicates that high scores on one variable tend to be accompanied by high scores on the other, and low scores on one variable by low scores on the other. A high negative correlation indicates that high scores on one variable tend to be accompanied by low scores on the other, and vice versa. A correlation of +1 represents perfect positive relationship and a correlation of -1 represents perfect negative relationship. A correlation of zero indicates that there is no relationship between scores on the two variables; knowing a person's score on one variable gives no clue as to what his score is on the other.

the experimental battery, make minor revisions in many of the others, and more substantial revisions in a few.

It was also decided to allot five minutes apiece for each of two very brief themes to be written by the students. These themes were to be on the topics, "What high school means to me", and "My views about an ideal occupation".

The four tests that were eliminated were:

Active Vocabulary
Verbal Reasoning
Scale Reading
Following Directions

The decision to eliminate these tests was made regretfully, but practical considerations dictated it. Although there seemed to be a good possibility that some of the four tests might make significant unique contributions to the battery, there was a limit to how many tests, even comparatively short ones, could fit into the available testing time.

The composition of the final battery is summarized in Table IV-2.

After the general decisions had been made as to content of the final battery, one major decision still remained to be made before the tests could be revised, put together, and organized into a single integrated battery. This was the question of how the various parts of the battery were to be handled in the initial data-processing phase, after the testing had been completed. It seems advisable to anticipate Chapter VI at this point, by indicating that two different kinds of machines--a "scoring machine" and a "document reader"--were available on which the initial processing could be done. The scoring machine would be suitable where one specific response to each option was

designated the "right" response and the number of "right" responses a student made on a group of items, in other words his "score", was to be determined. The document reader would be suitable where it was anticipated that instead of combining items into scores it would be desirable to have a record of the specific responses to individual items, so that these could be analyzed separately, or perhaps combined in a variety of ways. Which way a particular test or group of items would be handled had to be decided long before the tests were given, since the two machines required different types of answer sheets. It was decided to fit the battery on five answer sheets, two of which would be processed on the scoring machine and the remaining three on the document reader. Document reader materials would include the Student Information Blank, the Interest Inventory, and basic identification data--student's name, address, sex, date of birth, grade in school, etc., while the Student Activities Inventory and most of the tests (all of them, in fact, except the Information Test) would be processed on the scoring machine. That left the Information Test as the only problem area. This test posed a problem because some of the items fell neatly into subject-matter areas where it would be appropriate to obtain scores immediately, while other items were of such a nature that it seemed appropriate to score them individually. The solution was a simple one. The test was split into two parts; Part I went on one of the answer sheets earmarked for the scoring machine, and Part II went on another answer sheet, of the type to be processed by the document reader.

The organization of the final battery into four half-day sessions, the tests appearing in each of the four booklets and on each of the five answer sheets, and the time allowances are all summarized in Table IV-3. Of the five answer sheets, designated A, B1, B2, C, and Record Form Z, Answer Sheets A and C are the two designed for the Scoring Machine, while Answer Sheets B1 and B2 and Record Form Z are the sheets for the Document Reader. For the benefit of those readers who would like to see what the answer sheets look like, Figures 1-6 show both sides of Answer Sheet C, Answer Sheet B2, and Record Form Z. Note the space allowed on both sides of Answer Sheet B2 for writing one of the five-minute themes (Question 395).

Dress Rehearsal

The very last step before freezing the battery into its final form was a full-scale "dress rehearsal", which took place in February 1960. The "stage" was Preston High School, in Preston, Maryland, where the full two days of the battery were administered. Since things went smoothly, only a few minor changes in time limits were necessary, before the presses could start rolling to turn out materials for testing a half-million students.

ANSWER SHEET B2 (SIDE 1)

TESTING NUMBER____ TODAY'S DATE Mo.____ Day____ Year____

NAME____ Last____ First____ Middle____ ☐ BOY ☐ GIRL

DATE OF BIRTH Mo.____ Day____ Yr.____

SCHOOL____

CITY____ STATE____

DO NOT WRITE ABOVE THIS LINE

Question 395
START HERE

Turn answer sheet and continue.

DO NOT WRITE IN OR BELOW THIS SPACE

PART I cont.

A B C D E F

51 ○ ○ ○ ○ ○ ○
52 ○ ○ ○ ○ ○ ○
53 ○ ○ ○ ○ ○ ○
54 ○ ○ ○ ○ ○ ○
55 ○ ○ ○ ○ ○ ○
56 ○ ○ ○ ○ ○ ○
57 ○ ○ ○ ○ ○ ○
58 ○ ○ ○ ○ ○ ○
59 ○ ○ ○ ○ ○ ○
60 ○ ○ ○ ○ ○ ○
61 ○ ○ ○ ○ ○ ○
62 ○ ○ ○ ○ ○ ○
63 ○ ○ ○ ○ ○ ○
64 ○ ○ ○ ○ ○ ○
65 ○ ○ ○ ○ ○
66 ○ ○ ○ ○ ○
67 ○ ○ ○ ○ ○
68 ○ ○ ○ ○ ○
69 ○ ○ ○ ○ ○
70 ○ ○ ○ ○ ○
71 ○ ○ ○ ○ ○
72 ○ ○ ○ ○ ○
73 ○ ○ ○ ○ ○
74 ○ ○ ○ ○ ○
75 ○ ○ ○ ○ ○
76 ○ ○ ○ ○ ○
77 ○ ○ ○ ○ ○
78 ○ ○ ○ ○ ○
79 ○ ○ ○ ○ ○
80 ○ ○ ○ ○ ○
81 ○ ○ ○ ○ ○
82 ○ ○ ○ ○ ○
83 ○ ○ ○ ○ ○

A B C D E F

MAKE NO STRAY MARKS
IN THIS AREA

A B C D E F

84 ○ ○ ○ ○ ○
85 ○ ○ ○ ○ ○
86 ○ ○ ○ ○ ○
87 ○ ○ ○ ○ ○
88 ○ ○ ○ ○ ○
89 ○ ○ ○ ○ ○
90 ○ ○ ○ ○ ○
91 ○ ○ ○ ○ ○ ○
92 ○ ○ ○ ○ ○ ○
93 ○ ○ ○ ○ ○ ○
94 ○ ○ ○ ○ ○ ○
95 ○ ○ ○ ○ ○ ○
96 ○ ○ ○ ○ ○ ○
97 ○ ○ ○ ○ ○ ○
98 ○ ○ ○ ○ ○ ○
99 ○ ○ ○ ○ ○ ○
100 ○ ○ ○ ○ ○ ○
101 ○ ○ ○ ○ ○ ○
102 ○ ○ ○ ○ ○ ○
103 ○ ○ ○ ○ ○ ○
104 ○ ○ ○ ○ ○ ○
105 ○ ○ ○ ○ ○ ○
106 ○ ○ ○ ○ ○ ○
107 ○ ○ ○ ○ ○ ○
108 ○ ○ ○ ○ ○ ○
109 ○ ○ ○ ○ ○ ○
110 ○ ○ ○ ○ ○ ○
111 ○ ○ ○ ○ ○ ○
112 ○ ○ ○ ○ ○ ○
113 ○ ○ ○ ○ ○ ○
114 ○ ○ ○ ○ ○ ○
115 ○ ○ ○ ○ ○ ○
116 ○ ○ ○ ○ ○ ○
117 ○ ○ ○ ○ ○ ○
118 ○ ○ ○ ○ ○ ○
119 ○ ○ ○ ○ ○ ○
120 ○ ○ ○ ○ ○ ○
121 ○ ○ ○ ○ ○ ○
122 ○ ○ ○ ○ ○ ○
123 ○ ○ ○ ○ ○ ○
124 ○ ○ ○ ○ ○ ○
125 ○ ○ ○ ○ ○ ○
126 ○ ○ ○ ○ ○ ○
127 ○ ○ ○ ○ ○ ○
128 ○ ○ ○ ○ ○ ○
129 ○ ○ ○ ○ ○ ○

PART II

A B C D E F

130 ○ ○ ○ ○ ○ ○
131 ○ ○ ○ ○
132 ○ ○ ○ ○ ○ ○
133 ○ ○ ○ ○ ○
134 ○ ○ ○ ○ ○ ○
135 ○ ○ ○ ○ ○ ○
136 ○ ○ ○ ○ ○
137 ○ ○ ○ ○ ○
138 ○ ○ ○ ○ ○ ○
139 ○ ○ ○ ○ ○ ○
140 ○ ○ ○ ○ ○ ○
141 ○ ○ ○ ○ ○ ○
142 ○ ○ ○ ○ ○ ○
143 ○ ○ ○ ○ ○ ○
144 ○ ○ ○ ○ ○ ○

A B C D E F

A B C D E F

145 ○ ○ ○ ○ ○ ○
146 ○ ○ ○ ○ ○ ○
147 ○ ○ ○ ○ ○ ○
148 ○ ○ ○ ○ ○ ○
149 ○ ○ ○ ○ ○ ○
150 ○ ○ ○ ○ ○ ○
151 ○ ○ ○ ○ ○ ○
152 ○ ○ ○ ○ ○ ○
153 ○ ○ ○ ○ ○ ○
154 ○ ○ ○ ○ ○ ○
155 ○ ○ ○ ○ ○ ○
156 ○ ○ ○ ○ ○ ○
157 ○ ○ ○ ○ ○ ○
158 ○ ○ ○ ○ ○ ○
159 ○ ○ ○ ○ ○ ○
160 ○ ○ ○ ○ ○ ○
161 ○ ○ ○ ○ ○ ○
162 ○ ○ ○ ○ ○ ○
163 ○ ○ ○ ○ ○ ○
164 ○ ○ ○ ○ ○ ○
165 ○ ○ ○ ○ ○ ○
166 ○ ○ ○ ○ ○ ○
167 ○ ○ ○ ○ ○ ○
168 ○ ○ ○ ○ ○ ○
169 ○ ○ ○ ○ ○ ○
170 ○ ○ ○ ○ ○ ○
171 ○ ○ ○ ○ ○ ○
172 ○ ○ ○ ○ ○ ○
173 ○ ○ ○ ○ ○ ○
174 ○ ○ ○ ○ ○ ○
175 ○ ○ ○ ○ ○ ○

A B C D E F

176 ○ ○ ○ ○ ○ ○
177 ○ ○ ○ ○ ○ ○
178 ○ ○ ○ ○ ○ ○
179 ○ ○ ○ ○ ○ ○
180 ○ ○ ○ ○ ○ ○
181 ○ ○ ○ ○ ○ ○
182 ○ ○ ○ ○ ○ ○
183 ○ ○ ○ ○ ○ ○
184 ○ ○ ○ ○ ○ ○
185 ○ ○ ○ ○ ○ ○
186 ○ ○ ○ ○ ○ ○
187 ○ ○ ○ ○ ○ ○
188 ○ ○ ○ ○
189 ○ ○ ○ ○ ○
190 ○ ○ ○ ○ ○ ○
191 ○ ○ ○ ○ ○
192 ○ ○ ○ ○ ○ ○
193 ○ ○ ○ ○ ○ ○
194 ○ ○ ○ ○ ○ ○
195 ○ ○ ○ ○
196 ○ ○ ○ ○ ○ ○
197 ○ ○ ○ ○ ○ ○
198 ○ ○ ○ ○ ○ ○
199 ○ ○ ○ ○ ○ ○
200 ○ ○ ○ ○ ○
201 ○ ○ ○ ○ ○ ○
202 ○ ○ ○ ○ ○ ○
203 ○ ○ ○ ○ ○ ○
204 ○ ○ ○ ○ ○ ○
205 ○ ○ ○ ○ ○ ○

A B C D E F

TESTING NUMBER

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨
⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨

FIRST FOUR LETTERS OF LAST NAME

⊕ Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ Ⓕ Ⓖ Ⓗ Ⓘ Ⓙ Ⓚ Ⓛ Ⓜ Ⓝ Ⓞ Ⓟ Ⓠ Ⓡ Ⓢ Ⓣ Ⓤ Ⓥ Ⓦ Ⓧ Ⓨ Ⓩ
⊕ Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ Ⓕ Ⓖ Ⓗ Ⓘ Ⓙ Ⓚ Ⓛ Ⓜ Ⓝ Ⓞ Ⓟ Ⓠ Ⓡ Ⓢ Ⓣ Ⓤ Ⓥ Ⓦ Ⓧ Ⓨ Ⓩ
⊕ Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ Ⓕ Ⓖ Ⓗ Ⓘ Ⓙ Ⓚ Ⓛ Ⓜ Ⓝ Ⓞ Ⓟ Ⓠ Ⓡ Ⓢ Ⓣ Ⓤ Ⓥ Ⓦ Ⓧ Ⓨ Ⓩ
⊕ Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ Ⓕ Ⓖ Ⓗ Ⓘ Ⓙ Ⓚ Ⓛ Ⓜ Ⓝ Ⓞ Ⓟ Ⓠ Ⓡ Ⓢ Ⓣ Ⓤ Ⓥ Ⓦ Ⓧ Ⓨ Ⓩ

PLEASE . . . keep your answer sheet clean.
Do not fold it or bend the corners.

TALENT TESTING PROGRAM **Student Information Blank**

Copyright, 1960, by University of Pittsburgh
All rights reserved. Printed in the U.S.A.

Printed by MRC, Inc. Iowa City, Iowa. Form No. 1041

This should be the end of your answer to question 395.

DO NOT WRITE ABOVE THIS LINE

Question 395 (Continued)

Finish in column at right

DO NOT WRITE IN OR BELOW THIS SPACE

PLEASE...keep your answer sheet clean.

Do not fold it or bend the corners.

PART III

A B C D E F G H I J K L — M N O P Q R S T U V W X Y Z — A B C D E F G H I J

206 ○○○○○○○○○○○○ 206 ○○○○○
207 ○○○○○○○○○○○○ 207 ○○○○○○○○○○○○○○ 207 ○○○○○○○○○
208 ○○○○○○○○○○○○ 208 ○○○○○
209 ○○○○○○○○○○○○ 209 ○○○○○○○○○○○○○○ 209 ○○○○○○○○
210 ○○○○○○○○○○○○ 210 ○○○○○○○○○○
211 ○○○○○○○○○○○○ 211 ○○○○○○○○○○○○○○ 211 ○○○○○○○○○○
212 ○○○○○○○○○○○○ 212 ○○○○○○○○○○○○○○ 212 ○○○○○○○○○○

PART IV

A B C D E F G H I J K L

213 ○○○○○○○○○○
214 ○○○○○○○○○○
215 ○○○○○○○○○○
216 ○○○○○○○○○○
217 ○○○○○○○○
218 ○○○○○○○○○○○
219 ○○○○○○○○○○○
220 ○○○○○○○○○○
221 ○○○○○○○○○○○○
222 ○○○○○○○○○○○○
223 ○○○○○○○○○○○○
224 ○○○○○○○○○○○○
225 ○○○○○○○○○○○○
226 ○○○○○○○○○○○○
227 ○○○○○○○○○○
228 ○○○○○○○○○○○○
229 ○○○○○○○○○○
230 ○○○○○○○○○○○○
231 ○○○○○○○○○○○○
232 ○○○○○○○○○○○○
233 ○○○○○○○○○○○○
234 ○○○○○○○○○○○○
235 ○○○○○○○○○○○○
236 ○○○○○○○○○○○○
237 ○○○○○○○○○
238 ○○○○○○○○○○○○
239 ○○○○○○○○○○○
240 ○○○○○○○○○○○

A B C D E F G H I J K L

PART V

A B C D E F

241 ○○○○○○
242 ○○○○○○
243 ○○○○○○
244 ○○○○○○
245 ○○○○○○
246 ○○○○○○
247 ○○○○○○
248 ○○
249 ○○
250 ○○
251 ○○
252 ○○
253 ○○
254 ○○
255 ○○
256 ○○
257 ○○
258 ○○
259 ○○
260 ○○
261 ○○
262 ○○
263 ○○
264 ○○
265 ○○
266 ○○
267 ○○
268 ○○
269 ○○
270 ○○
271 ○○
272 ○○

A B

MAKE NO STRAY MARKS IN THIS AREA

A B C D E F

273 ○○
274 ○○
275 ○○
276 ○○
277 ○○
278 ○○
279 ○○
280 ○○
281 ○○
282 ○○
283 ○○
284 ○○○○○○
285 ○○○○○○
286 ○○○○○○
287 ○○○○○○
288 ○○○○○○
289 ○○○○○○
290 ○○○○○○
291 ○○○○○○
292 ○○○○○○
293 ○○○○○○
294 ○○○○○○
295 ○○
296 ○○○○

PART VI

A B C D E F

297 ○○○○○
298 ○○○○○
299 ○○○○○
300 ○○○○○
301 ○○○○○
302 ○○○○
303 ○○○○○
304 ○○○○○○
305 ○○○
306 ○○
307 ○○○○○○
308 ○○○○○○
309 ○○○○○○
310 ○○○○○○
311 ○○○○○○
312 ○○○○○○
313 ○○○○○○
314 ○○○○○○
315 ○○○○○○
316 ○○○○○○
317 ○○○○○○
318 ○○○○○○
319 ○○○○○○
320 ○○○○○○
321 ○○○○○○
322 ○○○○○○
323 ○○○○○○
324 ○○○○○○
325 ○○○○○○
326 ○○○○○○
327 ○○○○○○
328 ○○○○○○
329 ○○○○○○
330 ○○○○○○
331 ○○○○○○
332 ○○○○○○
333 ○○○○○

A B C D E F

A B C D E F

334 ○○○○○
335 ○○○○○
336 ○○○○○
337 ○○○○○○
338 ○○○○○○
339 ○○○○○
340 ○○○○○○
341 ○○○○○○
342 ○○○○○○
343 ○○○○○○
344 ○○○○○○
345 ○○○○○○
346 ○○○○○○
347 ○○○○○○
348 ○○○○○○
349 ○○○○○○
350 ○○○○○○
351 ○○○○○○
352 ○○○○○○
353 ○○○○○○
354 ○○○○○○
355 ○○○○○○
356 ○○○○○○
357 ○○○○○○
358 ○○○○○○
359 ○○○○○○
360 ○○○○○○
361 ○○○○○○
362 ○○○○○○
363 ○○○○○
364 ○○○○○○
365 ○○○○○○
366 ○○○○○○
367 ○○○○○○
368 ○○○○○○
369 ○○○○○○
370 ○○○○○
371 ○○○○○
372 ○○○○○○
373 ○○○○○○
374 ○○○○○○

PART VII

A B C D E F

375 ○○○○○○
376 ○○○○○○
377 ○○○○○○
378 ○○○○○○
379 ○○○○○○
380 ○○○○○○
381 ○○○○○○
382 ○○○○○○
383 ○○○○○○
384 ○○○○○○
385 ○○○○○○
386 ○○○○○○
387 ○○○○○○
388 ○○○○○○
389 ○○○○○○
390 ○○○○○○
391 ○○○○○
392 ○○○○○
393 ○○○○○○
394 ○○○○○○

A B C D E F

ANSWER SHEET B2 (SIDE 2)

ANSWER SHEET C (SIDE 1) • TALENT TESTING PROGRAM

Arithmetic Computation

A B C D E
S1 ○○○○●

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	10 ○○○○○	19 ○○○○○	28 ○○○○○	37 ○○○○○	46 ○○○○○	55 ○○○○○	64 ○○○○○
2 ○○○○○	11 ○○○○○	20 ○○○○○	29 ○○○○○	38 ○○○○○	47 ○○○○○	56 ○○○○○	65 ○○○○○
3 ○○○○○	12 ○○○○○	21 ○○○○○	30 ○○○○○	39 ○○○○○	48 ○○○○○	57 ○○○○○	66 ○○○○○
4 ○○○○○	13 ○○○○○	22 ○○○○○	31 ○○○○○	40 ○○○○○	49 ○○○○○	58 ○○○○○	67 ○○○○○
5 ○○○○○	14 ○○○○○	23 ○○○○○	32 ○○○○○	41 ○○○○○	50 ○○○○○	59 ○○○○○	68 ○○○○○
6 ○○○○○	15 ○○○○○	24 ○○○○○	33 ○○○○○	42 ○○○○○	51 ○○○○○	60 ○○○○○	69 ○○○○○
7 ○○○○○	16 ○○○○○	25 ○○○○○	34 ○○○○○	43 ○○○○○	52 ○○○○○	61 ○○○○○	70 ○○○○○
8 ○○○○○	17 ○○○○○	26 ○○○○○	35 ○○○○○	44 ○○○○○	53 ○○○○○	62 ○○○○○	71 ○○○○○
9 ○○○○○	18 ○○○○○	27 ○○○○○	36 ○○○○○	45 ○○○○○	54 ○○○○○	63 ○○○○○	72 ○○○○○

Memory for Words

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	4 ○○○○○	7 ○○○○○	10 ○○○○○	13 ○○○○○	16 ○○○○○	19 ○○○○○	22 ○○○○○
2 ○○○○○	5 ○○○○○	8 ○○○○○	11 ○○○○○	14 ○○○○○	17 ○○○○○	20 ○○○○○	23 ○○○○○
3 ○○○○○	6 ○○○○○	9 ○○○○○	12 ○○○○○	15 ○○○○○	18 ○○○○○	21 ○○○○○	24 ○○○○○

Memory for Sentences

A E O L T
S1 ○○●○○

A I L P R
S2 ●○○○○

A E O T X	A L O R U	A C L O W	A E I L R	A E I O T	A E T U W	E L O P R	E I O L T
1 ○○○○○	3 ○○○○○	5 ○○○○○	7 ○○○○○	9 ○○○○○	11 ○○○○○	13 ○○○○○	15 ○○○○○
A E N S R	A H I L U	T A I R U	A H I L O	A E P R T	A E I L O	I O P R T	R I A X W
2 ○○○○○	4 ○○○○○	6 ○○○○○	8 ○○○○○	10 ○○○○○	12 ○○○○○	14 ○○○○○	16 ○○○○○

Mathematics Part I

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D	A B C D	A B C D E
1 ○○○○○	3 ○○○○○	5 ○○○○○	7 ○○○○○	9 ○○○○○	11 ○○○○	13 ○○○○	15 ○○○○○
2 ○○○○○	4 ○○○○○	6 ○○○○○	8 ○○○○○	10 ○○○○	12 ○○○○	14 ○○○○	16 ○○○○○

Mathematics Part II

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
17 ○○○○○	20 ○○○○○	23 ○○○○○	26 ○○○○○	29 ○○○○○	32 ○○○○○	35 ○○○○○	38 ○○○○○
18 ○○○○○	21 ○○○○○	24 ○○○○○	27 ○○○○○	30 ○○○○○	33 ○○○○○	36 ○○○○○	39 ○○○○○
19 ○○○○○	22 ○○○○○	25 ○○○○○	28 ○○○○○	31 ○○○○○	34 ○○○○○	37 ○○○○○	40 ○○○○○

Mathematics Part III

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
41 ○○○○○	43 ○○○○○	45 ○○○○○	47 ○○○○○	49 ○○○○○	51 ○○○○○	53 ○○○○○
42 ○○○○○	44 ○○○○○	46 ○○○○○	48 ○○○○○	50 ○○○○○	52 ○○○○○	54 ○○○○○

English Part I

A B C D E
S1 ●○○○○

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
							22 ○○○○
1 ○○○○○	4 ○○○	7 ○○○○○	10 ○○○○○	13 ○○○○○	16 ○○○○○	19 ○○○○○	23 ○○○○○
2 ○○○○	5 ○○○○○	8 ○○○○○	11 ○○○○	14 ○○○○○	17 ○○○○○	20 ○○○○	24 ○○○○
3 ○○○○○	6 ○○○○	9 ○○○○	12 ○○○○	15 ○○○○	18 ○○○○	21 ○○○○○	25 ○○○○○

English Part II

A B C D E	A B C D	A B C	A B C
26 ○○○○	29 ○○○	32 ○○○	35 ○○○
27 ○○○	30 ○○○○	33 ○○○	36 ○○○
28 ○○○○○	31 ○○○	34 ○○○	37 ○○○

English Part III

A B C D E
S1 ○●○
S2 ○○○○●

Section a A B C D E	A B C D E	A B C D E	0 1 2	0 1 2
38 ○○○○○	44 ○○○○○	51 ○○○○	56 ○○○	63 ○○○
39 ○○○○	45 ○○○	52 ○○○○○	57 ○○○	64 ○○○
40 ○○○○	46 ○○○○○	53 ○○○○○	58 ○○○	
41 ○○○	47 ○○○○○	Section b	59 ○○○	
42 ○○○○○	48 ○○○○○	0 1 2	60 ○○○	
43 ○○○	49 ○○○○○	54 ○○○	61 ○○○	
	50 ○○○○○	55 ○○○	62 ○○○	

English Part IV

A B C D E
S1 ○○●○○
S2 ○○○○●

A B C D E	A B C D E	A B C D E	A B C D E
65 ○○○○○	69 ○○○○○	73 ○○○○○	77 ○○○○○
66 ○○○○○	70 ○○○○○	74 ○○○○○	78 ○○○○○
67 ○○○○○	71 ○○○○○	75 ○○○○○	79 ○○○○○
68 ○○○○○	72 ○○○○○	76 ○○○○○	80 ○○○○○

English Part V

C S	C S	C S	C S	C S
81 ○ ○	88 ○ ○	95 ○ ○	102 ○ ○	109 ○ ○
82 ○ ○	89 ○ ○	96 ○ ○	103 ○ ○	110 ○ ○
83 ○ ○	90 ○ ○	97 ○ ○	104 ○ ○	111 ○ ○
84 ○ ○	91 ○ ○	98 ○ ○	105 ○ ○	112 ○ ○
85 ○ ○	92 ○ ○	99 ○ ○	106 ○ ○	113 ○ ○
86 ○ ○	93 ○ ○	100 ○ ○	107 ○ ○	114 ○ ○
87 ○ ○	94 ○ ○	101 ○ ○	108 ○ ○	

Abstract Reasoning

A B C D E
S1 ○○○○●
S2 ○○○○●
S3 ○○●○○

A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	5 ○○○○○	9 ○○○○○	13 ○○○○○
2 ○○○○○	6 ○○○○○	10 ○○○○○	14 ○○○○○
3 ○○○○○	7 ○○○○○	11 ○○○○○	15 ○○○○○
4 ○○○○○	8 ○○○○○	12 ○○○○○	

PLEASE . . .

Keep your answer sheet as clean as possible. Be careful not to fold the answer sheet or to bend the corners.

TESTING NUMBER	⓪①②③④⑤⑥⑦⑧⑨
	⓪①②③④⑤⑥⑦⑧⑨
	⓪①②③④⑤⑥⑦⑧⑨
	⓪①②③④⑤⑥⑦⑧⑨
	⓪①②③④⑤⑥⑦⑧⑨
	⓪①②③④⑤⑥⑦⑧⑨
MOTHER'S MAIDEN NAME	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
FATHER'S FIRST NAME	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
FIRST 8 LETTERS OF LAST NAME	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
MIDDLE INITIAL	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ
FIRST INITIAL	⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ

Copyright, 1960, by University of Pittsburgh

Printed in the U.S.A.

Printed by [illegible]

ANSWER SHEET C (SIDE 2)

YOUR TESTING No.________ TODAY'S DATE Mo. Day Yr.

YOUR NAME________ Last First Middle

SCHOOL________

CITY________ STATE________

Mechanical Reasoning

A B C
S1 ●○○

A B C D	A B C D E	A B C D E	A B C C E
1 ○○○○	6 ○○○○○	11 ○○○○○	16 ○○○○
2 ○○○○	7 ○○○	12 ○○○○	17 ○○○○
3 ○○○○	8 ○○○	13 ○○○	18 ○○○○○
4 ○○○	9 ○○○○○	14 ○○○	19 ○○○
5 ○○○	10 ○○○	15 ○○○○	20 ○○○○

Disguised Words

A B C D E
S1 ○●○○○

A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	9 ○○○○○	17 ○○○○○	25 ○○○○○
2 ○○○○○	10 ○○○○○	18 ○○○○○	26 ○○○○○
3 ○○○○○	11 ○○○○○	19 ○○○○○	27 ○○○○○
4 ○○○○○	12 ○○○○○	20 ○○○○○	28 ○○○○○
5 ○○○○○	13 ○○○○○	21 ○○○○○	29 ○○○○○
6 ○○○○○	14 ○○○○○	22 ○○○○○	30 ○○○○○
7 ○○○○○	15 ○○○○○	23 ○○○○○	
8 ○○○○○	16 ○○○○○	24 ○○○○○	

Creativity

A B C D E
S1 ○○●○○

A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	6 ○○○○○	11 ○○○○○	16 ○○○○○
2 ○○○○○	7 ○○○○○	12 ○○○○○	17 ○○○○○
3 ○○○○○	8 ○○○○○	13 ○○○○○	18 ○○○○○
4 ○○○○○	9 ○○○○○	14 ○○○○○	19 ○○○○○
5 ○○○○○	10 ○○○○○	15 ○○○○○	20 ○○○○○

Clerical Checking

S D
S1 ○ ●
S2 ● ○

S D	S D	S D	S D	S D	S D	S D
1 ○ ○	15 ○ ○	29 ○ ○	43 ○ ○	57 ○ ○	71 ○ ○	73 ○ ○
2 ○ ○	16 ○ ○	30 ○ ○	44 ○ ○	58 ○ ○	72 ○ ○	74 ○ ○
3 ○ ○	17 ○ ○	31 ○ ○	45 ○ ○	59 ○ ○		
4 ○ ○	18 ○ ○	32 ○ ○	46 ○ ○	60 ○ ○		
5 ○ ○	19 ○ ○	33 ○ ○	47 ○ ○	61 ○ ○		
6 ○ ○	20 ○ ○	34 ○ ○	48 ○ ○	62 ○ ○		
7 ○ ○	21 ○ ○	35 ○ ○	49 ○ ○	63 ○ ○		
8 ○ ○	22 ○ ○	36 ○ ○	50 ○ ○	64 ○ ○		
9 ○ ○	23 ○ ○	37 ○ ○	51 ○ ○	65 ○ ○		
10 ○ ○	24 ○ ○	38 ○ ○	52 ○ ○	66 ○ ○		
11 ○ ○	25 ○ ○	39 ○ ○	53 ○ ○	67 ○ ○		
12 ○ ○	26 ○ ○	40 ○ ○	54 ○ ○	68 ○ ○		
13 ○ ○	27 ○ ○	41 ○ ○	55 ○ ○	69 ○ ○		
14 ○ ○	28 ○ ○	42 ○ ○	56 ○ ○	70 ○ ○		

Visual'z'n in 2 Dimensions

A B C D E
S1 ○○○●○

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	4 ○○○○○	7 ○○○○○	10 ○○○○○	13 ○○○○○	16 ○○○○○	19 ○○○○○	22 ○○○○○
2 ○○○○○	5 ○○○○○	8 ○○○○○	11 ○○○○○	14 ○○○○○	17 ○○○○○	20 ○○○○○	23 ○○○○○
3 ○○○○○	6 ○○○○○	9 ○○○○○	12 ○○○○○	15 ○○○○○	18 ○○○○○	21 ○○○○○	24 ○○○○○

Reading Comprehension

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	7 ○○○○○	13 ○○○○○	19 ○○○○○	25 ○○○○○	31 ○○○○○	37 ○○○○○	43 ○○○○○
2 ○○○○○	8 ○○○○○	14 ○○○○○	20 ○○○○○	26 ○○○○○	32 ○○○○○	38 ○○○○○	44 ○○○○○
3 ○○○○○	9 ○○○○○	15 ○○○○○	21 ○○○○○	27 ○○○○○	33 ○○○○○	39 ○○○○○	45 ○○○○○
4 ○○○○○	10 ○○○○○	16 ○○○○○	22 ○○○○○	28 ○○○○○	34 ○○○○○	40 ○○○○○	46 ○○○○○
5 ○○○○○	11 ○○○○○	17 ○○○○○	23 ○○○○○	29 ○○○○○	35 ○○○○○	41 ○○○○○	47 ○○○○○
6 ○○○○○	12 ○○○○○	18 ○○○○○	24 ○○○○○	30 ○○○○○	36 ○○○○○	42 ○○○○○	48 ○○○○○

Visual'z'n in 3 Dimensions

A B C D E
S1 ○○○○●

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	3 ○○○○○	5 ○○○○○	7 ○○○○○	9 ○○○○○	11 ○○○○○	13 ○○○○○	15 ○○○○○
2 ○○○○○	4 ○○○○○	6 ○○○○○	8 ○○○○○	10 ○○○○○	12 ○○○○○	14 ○○○○○	16 ○○○○○

Word Functions

A B C D E
S1 ○○●○○
S2 ○○●○○

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	4 ○○○○○	7 ○○○○○	10 ○○○○○	13 ○○○○○	16 ○○○○○	19 ○○○○○	22 ○○○○○
2 ○○○○○	5 ○○○○○	8 ○○○○○	11 ○○○○○	14 ○○○○○	17 ○○○○○	20 ○○○○○	23 ○○○○○
3 ○○○○○	6 ○○○○○	9 ○○○○○	12 ○○○○○	15 ○○○○○	18 ○○○○○	21 ○○○○○	24 ○○○○○

Table Reading

A B C D E
S1 ○○●○○
S2 ○○○○●

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	10 ○○○○○	19 ○○○○○	28 ○○○○○	37 ○○○○○	46 ○○○○○	55 ○○○○○	64 ○○○○○
2 ○○○○○	11 ○○○○○	20 ○○○○○	29 ○○○○○	38 ○○○○○	47 ○○○○○	56 ○○○○○	65 ○○○○○
3 ○○○○○	12 ○○○○○	21 ○○○○○	30 ○○○○○	39 ○○○○○	48 ○○○○○	57 ○○○○○	66 ○○○○○
4 ○○○○○	13 ○○○○○	22 ○○○○○	31 ○○○○○	40 ○○○○○	49 ○○○○○	58 ○○○○○	67 ○○○○○
5 ○○○○○	14 ○○○○○	23 ○○○○○	32 ○○○○○	41 ○○○○○	50 ○○○○○	59 ○○○○○	68 ○○○○○
6 ○○○○○	15 ○○○○○	24 ○○○○○	33 ○○○○○	42 ○○○○○	51 ○○○○○	60 ○○○○○	69 ○○○○○
7 ○○○○○	16 ○○○○○	25 ○○○○○	34 ○○○○○	43 ○○○○○	52 ○○○○○	61 ○○○○○	70 ○○○○○
8 ○○○○○	17 ○○○○○	26 ○○○○○	35 ○○○○○	44 ○○○○○	53 ○○○○○	62 ○○○○○	71 ○○○○○
9 ○○○○○	18 ○○○○○	27 ○○○○○	36 ○○○○○	45 ○○○○○	54 ○○○○○	63 ○○○○○	72 ○○○○○

Object Inspection

A B C D E
S1 ○○●○○

A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E
1 ○○○○○	6 ○○○○○	11 ○○○○○	16 ○○○○○	21 ○○○○○	26 ○○○○○	31 ○○○○○	36 ○○○○○
2 ○○○○○	7 ○○○○○	12 ○○○○○	17 ○○○○○	22 ○○○○○	27 ○○○○○	32 ○○○○○	37 ○○○○○
3 ○○○○○	8 ○○○○○	13 ○○○○○	18 ○○○○○	23 ○○○○○	28 ○○○○○	33 ○○○○○	38 ○○○○○
4 ○○○○○	9 ○○○○○	14 ○○○○○	19 ○○○○○	24 ○○○○○	29 ○○○○○	34 ○○○○○	39 ○○○○○
5 ○○○○○	10 ○○○○○	15 ○○○○○	20 ○○○○○	25 ○○○○○	30 ○○○○○	35 ○○○○○	40 ○○○○○

PLEASE PRINT: LAST NAME________ FIRST NAME________ MIDDLE NAME________

SCHOOL________ CITY________ STATE________

YOUR TESTING NUMBER________ TODAY'S DATE________ Mo. Day Year

DATE OF BIRTH________ Mo. Day Yr. AGE________

Are you a student in the school in which you are taking these tests?

- ☐ Yes.
- ☐ No, I am not a student in any school.
- ☐ No, I am a student in the following school:

School________

City________ State________

Mark whether you are a boy or girl ☐ Boy ☐ Girl

MAKE NO STRAY MARKS BELOW THIS LINE

YOUR TESTING NUMBER

⓪⓪⓪⓪⓪⓪
①①①①①①
②②②②②②
③③③③③③
④④④④④④
⑤⑤⑤⑤⑤⑤
⑥⑥⑥⑥⑥⑥
⑦⑦⑦⑦⑦⑦
⑧⑧⑧⑧⑧⑧
⑨⑨⑨⑨⑨⑨

Fill in spaces above. Then copy in boxes below; one letter or digit per box. Leave a blank space between name for instance the name Mary Ruth Jones would be written like this:

JONES MARY RUTH

YOUR NAME

Last name, First name, Middle name. Leave one space between names.

⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕⊕
ⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶⒶ
ⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷⒷ
ⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸⒸ
ⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹⒹ
ⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺⒺ
ⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻⒻ
ⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼⒼ
ⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽⒽ
ⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾⒾ
ⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿⒿ
ⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀⓀ
ⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁⓁ
ⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂⓂ
ⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃⓃ
ⓄⓄⓄⓄⓄⓄⓄⓄⓄⓄⓄⓄⓄⓄ ⓄⓄⓄⓄⓄⓄ
ⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅⓅ
ⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆⓆ
ⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇⓇ
ⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈⓈ
ⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉⓉ
ⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊⓊ
ⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋⓋ
ⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌⓌ
ⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍⓍ
ⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎⓎ
ⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏⓏ

DATE OF BIRTH Mo	DATE OF BIRTH Day	DATE OF BIRTH Yr	AGE AT LAST BIRTH-DAY
	1		
	2		
	3		
	4		
	5		
	6	Bef'	
	7	'38	8
	8	'39	9
	9	'40	10
JAN	10	'41	11
FEB	11	'42	12
MAR	12	'43	13
APR	13	'44	14
MAY	14	'45	15
JUN	15	'46	16
JUL	16	'47	17
AUG	17	'48	18
SEP	18	'49	19
OCT	19	'50	20
NOV	20	'51	21
DEC	21	'52	
	22		
	23		
	24		
	25		If Older Than 21, Mark Below ☐
	26		
	27		
	28		
	29		
	30		
	31		

Blacken just one space in each of these two columns.

(the bottom space if still in school.)

GRADE YOU ARE NOW IN		LAST GRADE COMPLETED	
☐	Grade 1	1st Half	☐
☐	" 1	2nd "	☐
☐	" 2	1st "	☐
☐	" 2	2nd "	☐
☐	" 3	1st "	☐
☐	" 3	2nd "	☐
☐	" 4	1st "	☐
☐	" 4	2nd "	☐
☐	" 5	1st "	☐
☐	" 5	2nd "	☐
☐	" 6	1st "	☐
☐	" 6	2nd "	☐
☐	Grade 7	1st Half	☐
☐	" 7	2nd "	☐
☐	" 8	1st "	☐
☐	" 8	2nd "	☐
☐	" 9	1st "	☐
☐	" 9	2nd "	☐
☐	" 10	1st "	☐
☐	" 10	2nd "	☐
☐	" 11	1st "	☐
☐	" 11	2nd "	☐
☐	" 12	1st "	☐
☐	" 12	2nd "	☐
☐	College Freshman	1st Half	☐
☐	" "	2nd "	☐
☐	" Sophomore	1st "	☐
☐	" "	2nd "	☐
☐	" Junior	1st "	☐
☐	" "	2nd "	☐
☐	" Senior	1st "	☐
☐	" "	2nd "	☐
☐	Ungraded Class		
☐	←Other→		☐
☐	I am no longer in school	I am still in school	☐

○○⓪①②③④⑤⑥⑦⑧⑨ ⊕ⒶⒷⒸⒹⒺⒻⒼⒽⒾⒿⓀⓁⓂⓃⓄⓅⓆⓇⓈⓉⓊⓋⓌⓍⓎⓏ

DO NOT MARK IN THESE SPACES

MAKE NO STRAY MARKS BELOW THIS LINE

Blacken just one of the four boxes below.

I. If you are a student in the school in which you are taking these tests, mark "Boy" or "Girl" here: → Boy ☐ 1 Girl ☐ 2

II. If you are not a student in the school in which you are taking these tests, mark "Boy" or "Girl" here → Boy ☐ 3 Girl ☐ 4

TALENT TESTING PROGRAM

RECORD FORM Z

(SIDE 1)

PLEASE . . . keep this record form clean.

Do not fold it or bend the corners.

Copyright, 1960, by University of Pittsburgh
All rights reserved. Printed in the U.S.A.

PRINT YOUR HOME ADDRESS BELOW:

ADDRESS ____________________

CITY ____________________ POSTAL ZONE ____________

STATE ____________________

Now copy in the boxes below; one letter or digit per box.
Leave a blank space between words.

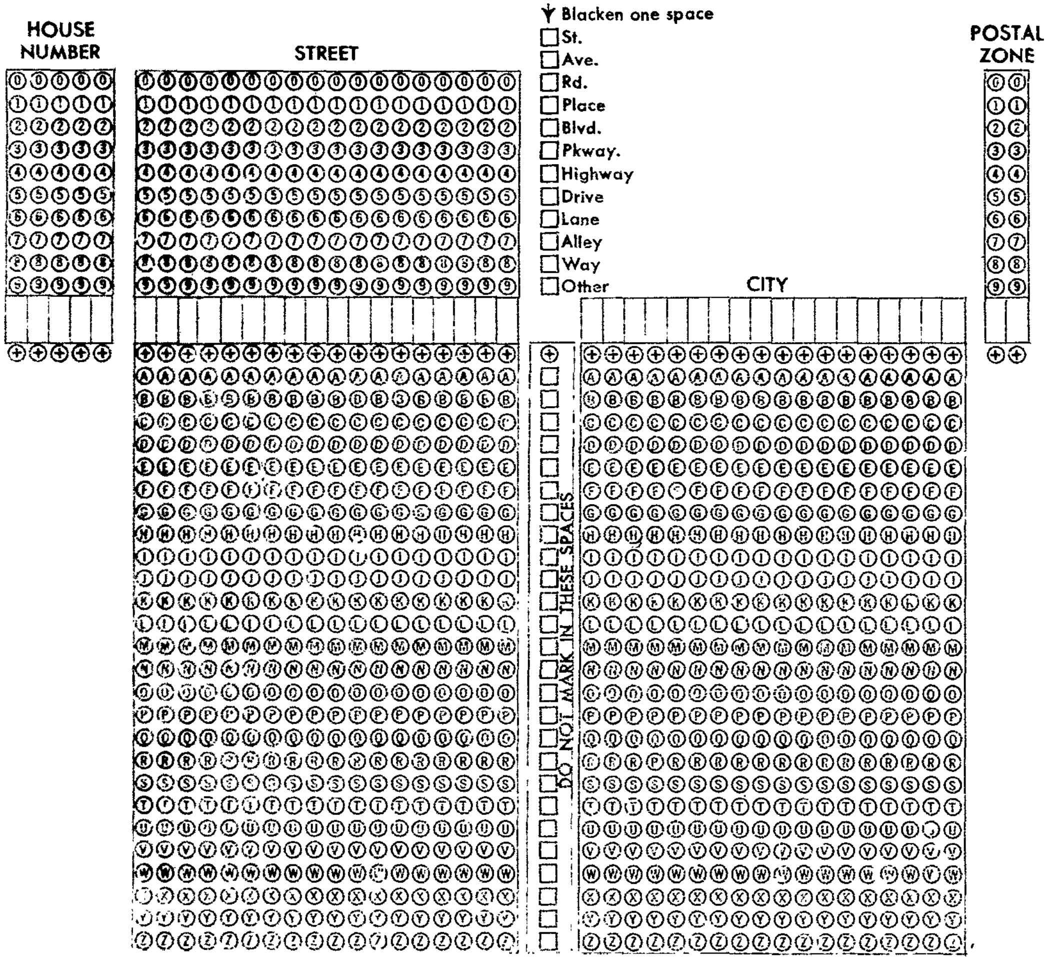
For instance: LOS ANGELES

PLEASE...keep this record form clean.

Do not fold it or bend the corners.

MAKE NO STRAY MARKS BELOW THIS LINE

YOUR HOME ADDRESS



STATE

Blacken one space

	⊕ ⊕	
Alabama	○ ○	Montana
Alaska	○ ○	Nebraska
Arizona	○ ○	Nevada
Arkansas	○ ○	N.H.
Calif.	○ ○	N.J.
Colorado	○ ○	N.M.
Conn.	○ ○	N.Y.
Delaware	○ ○	N.C.
D.C.	○ ○	N.D.
Florida	○ ○	Ohio
Georgia	○ ○	Oklahoma
Hawaii	○ ○	Oregon
Idaho	○ ○	Pa.
Illinois	○ ○	Rhode Is.
Indiana	○ ○	S.C.
Iowa	○ ○	S.D.
Kansas	○ ○	Tenn.
Kentucky	○ ○	Texas
La.	○ ○	Utah
Maine	○ ○	Vermont
Maryland	○ ○	Virginia
Mass.	○ ○	Wash.
Michigan	○ ○	W.Va.
Minn.	○ ○	Wisconsin
Miss.	○ ○	Wyoming
Missouri	○ ○	Other

DO NOT

WRITE

IN THIS

SPACE

RECORD FORM Z

(SIDE 2)

CHAPTER V

HOW THE TESTS WERE GIVEN

At 9:00 a.m. on March 1, 1960, all 225 students of the Riverside Indian School in Anadarko, Oklahoma were seated in their homerooms. Today, they were to take some tests, but no one had studied! Furthermore, all of the nine teachers knew no one had studied. No preparation was possible for the Talent tests, since they covered all school subjects--and some not taught in school. Instead of the students studying, the teachers had prepared for the test! They had practically "gone to school" just to learn how to give these tests. They had attended lectures, conferences, and read a book of instructions just to get ready for this day.

On each teacher's desk were the test books for that morning. But, in addition, there were Answer Sheets A, Record Forms Z, Student Identification Cards, a Test Record Form, extra pencils, a stopwatch, and, of course the instruction book.

Written on each blackboard was:

DO NOT TURN OR OPEN YOUR TEST
BOOKLETS UNTIL TOLD TO DO SO.

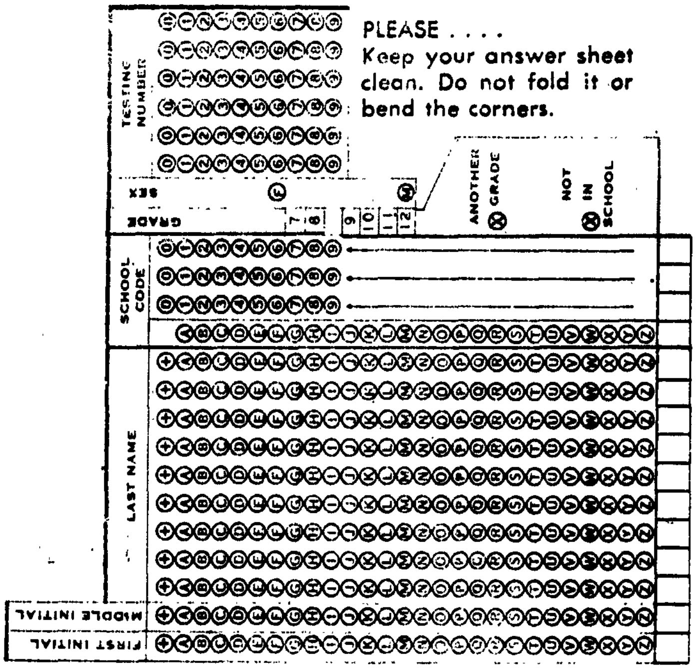
Test Booklets A were passed out, then Answer Sheets A, and finally Student Identification Cards.

When all the materials had been passed out, each teacher read this statement:

> "Starting today you will be taking part in a nation-wide study of high school talents. As a student in one of the high schools selected, you are part of a group that represents ALL high schools throughout the country. Only ONE in every twenty high schools will take part in this project.
>
> "One of the things this study is trying to find out is what your best talents are! The tests will cover a wide range of topics. Some of the questions will be easy and some will be hard, but don't be discouraged by the hard ones, because no one is expected to be equally good in all areas. Just work as quickly and accurately as you can and try your very best.
>
> "Our school will receive a summary of your test results. This information can be useful to your teachers, counselors, and advisors in helping you to make plans for after high school."

For the next half hour, teachers read instructions on how to mark answer sheets and record forms. The students

had to do more than just write their names. Since the answer sheets were devised so that they could be "read" by a machine, students "wrote" their names by blackening in spaces to spell out their names, school code, grade, sex, and testing number. Here is what this part of the answer sheet looked like before it was filled in:



PLEASE . . . .
Keep your answer sheet clean. Do not fold it or bend the corners.

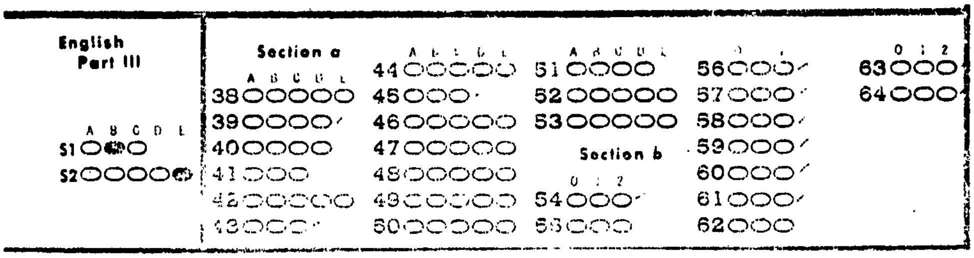
Later on, the scoring machine would be able to "read" the information about each student while scoring his answer sheet!

At about 9:30 a.m., the students had been read the instructions for the first test — the Information Test. They opened their test books and began answering questions on literature, mechanics, mathematics, food, physics, music, electricity, biology, chemistry, home economics, geography, zoology, social studies, farming, space travel, sports, and many other topics.

The students felt that this was a good test on which to start. If a boy didn't know much about cooking, he surely

could answer the questions about fishing or baseball or monkey wrenches. Everyone could answer some questions on some subject!

For two days, the 225 students of the Riverside Indian School answered questions by blackening in spaces on answer sheets. For each question, the student chose an answer from among several which were offered, and then blackened in that space on an answer sheet. Part of an answer sheet is shown below:



At the end of the two days, each student had been asked nearly 2000 questions-- not only what he knew, but what he was interested in, what kind of a person he was, what his plans were, and many other questions. Each student probably answered more questions and told more about himself than he ever had before, or ever would again!

At 3:57 p.m. on March 2, the testing was over. America's first children at the Riverside Indian School were the first to finish Project Talent with one hundred percent of the students taking part.

The scene above took place in 1356 other schools from Madawska, Maine to Wailuku, Hawaii in the spring of 1960. Although not every school finished with 100% participation, schools all over the country reported that attendence was higher on the testing days. Students who took all of the tests became members of the Project. They were issued

membership cards, each with his name and individual testing number. Years from now, their card will still identify them as having participated in this history making project.

The People Who Arranged to Give the Tests

In order to arrange a study of the high school students of our country, thousands of people had to cooperate. Over 18,000 teachers actually gave the tests in the schools. Before they began, they spent many hours of their personal time studying over the procedures for actually giving the tests. They also attended faculty meetings in which the exact procedures were reviewed in detail.

In addition to all these many teachers who gave the tests, school principals, guidance people, and superintendents of schools spent many hours in helping to prepare the teachers to give the tests.

The responsibility for deciding to give the tests was that of the superintendents and principals. Occasionally, Board of Education members were brought in to consider the problem. The decision to participate was made after these people gave careful consideration of the relative merits to be gained from giving these tests, versus the knowledge that might be gained by two regular days of instruction in class rooms.

The superintendents, Board Members, principals, teachers, and guidance people gave time and effort without any additional salary, or with little special recognition.

Almost a year before the testing took place, the first steps were taken which would eventually get the tests on the desks of the students.

The Regional Coordinators

The Story of Project Talent was taken directly to the school people and to the students. Only through their complete understanding and cooperation could the Project goals be realized. This job was done by 90 Regional Coordinators.

In the summer of 1959, the Project Staff began assembling a list of outstanding educators and psychologists who could be expected to be able and willing to work with the schools in administering the tests. By late fall, the Regional Coordinators had been selected.

These men and women were outstanding not only in their professions, but had long records of success in dealing with other people. Many of the Regional Coordinators accepted the job, knowing that it would have to be done on time outside of their regularly held positions. Here are some of the positions they held:

Head of a Department (Education, Psychology, or Guidance) of a University

Professor of Education, Psychology or Guidance

Dean of a Graduate School

Director of Research of a Public School System

State Guidance Supervisor

Director of a High School

Director of a Student Counseling Bureau

Dean of a University School of Education

Director of a Guidance or Counseling Center

Assistant Superintendent of Schools

Director of Bureau of Special Services for Pupils

Director of Testing and Evaluation Services for Board of Education

Director, Bureau of Educational Research and Service

State Director of Testing, Guidance and Counseling

Coordinator of Secondary School of Education

University Examiner

Director of Student Personnel (University)

and many others.

Some of the Regional Coordinators were in private practice; others had retired, but wanted to help in the project.

One link among the Regional Coordinators was their common interest in the young people of our country, in their education, guidance, and vocational success. They have devoted their professional lives persistently toward understanding, guiding, counseling and educating students, and toward improving education in the United States.

As it turned out, they were called upon to demonstrate some personal qualities which were not expected of them.

Regional Coordinators Visit the Schools

The Regional Coordinators made trips to each of the schools to tell the story of the Project to school people and to students, and to help in preparing the teachers to give the tests.

Getting an appointment was sometimes challenging. One coordinator reported:

> "One school didn't have a phone. In fact, the principal didn't have a phone in his house, either."

Then, getting to a school was often more challenging. One coordinator reported:

> "Some of my travel was done at the very time when heavy winter snows made some of the mountain passes impassable. I had to buy chains to reach one city."

From another:

> "For three successive weeks I had to cancel a trip to the bulk of my schools because of snowstorms that made it questionable whether I would have gotten through. As it developed, one or more of the schools had dismissed classes on each of those three weeks, so the trip would have been fruitless in any event."

And from another:

> "I located the town on the map, and found a road that cut out quite a bit of the distance. My car got mired down in the road. . .tow truck. . .I ended up driving 200 miles around to get to a place that was only 25 miles away."

Some of the reports from the Regional Coordinators describe best the general reception they received at the schools:

> ". . . . . . .the eagerness of many school people to participate in a project of national scope and significance. Some principals and superintendents hardly let me finish my spiel before they volunteered the help of their schools. Although I always represented the school selection as a random one, there was a strong tendency on the part of school people to give participation in the project some honorific value. . ."

> ". . .it was most gratifying to receive the full and unselfish cooperation of the schools . . . .Not once did I encounter any outward animosity, distrust, or an indifferent attitude toward this project. Every school, large or small, gave the impression that they are sincerely interested in doing all that they possibly can for the students in their schools, and were delighted at the opportunity to participate in such a project even though it demanded they interrupt their regular school schedule for

> two full days. . .The people in these schools are deeply concerned with the educative process and are interested in doing what is best within the limitations under which they operate for their students."
>
> ". . . . .probably my most vivid impression would be the excellent cooperation of the secondary school personnel involved. . . . ."
>
> ". . . .principals at at least two schools were suspicious of me--thinking that their schools had been selected because of their poverty--both instructionally and economically. . .Interestingly enough, though, I encountered no real hostility anywhere."

These are, of course, subjective and impressionistic accounts. They could be affected by the time of day or the day of the week on which the school was visited. They do, however, suggest that school people generally were eager to help in a project which would benefit students now and in the future.

Preparing the Teachers to Do the Testing

Giving a Standardized Test.

What is so different about a teacher giving a test on the subject he is teaching, and giving standardized tests such as the Project Talent tests?

Teachers give tests regularly as a part of course work so that they can check on how well the instruction is going over, if the students are progressing at the pace they should, if they are studying and for a variety of other reasons. Very often the content of these tests reviews the material which has been taught in a previous week or month. The teacher can check on how well the students are studying their assignments and, in addition, have some basis for assigning grades in courses. Generally, each teacher spends considerable time outside of class making up these tests, and then grading them. A test may take a class period, or it

may be a quiz which takes only 5 minutes. The length and content depend largely upon the individual teacher.

A standardized test, however, is constructed so that the questions will measure some specific kind of knowledge, trait, interest or other facet of an individual. For example, in the subject matter area of mechanics, there are certain basic ideas or principles which are important for everything in the field, such as gears, pulleys, levers, wheels, and others. A test of mechanical aptitude, then, would include questions on these principles. The total score on the test, thus, indicates how well the student understands the principles which are involved in mechanics. If a test is well constructed, a student might take it again and still get practically the same score.

Proceeding further, a standardized test is always administered the same way to all students who take it. The direction for the test are written down in a manual, and are read to the students. Generally, there is a time limit for a standardized test, which must be carefully kept, since if one student were allowed a few more seconds than the standard or established time, he would have an advantage over other students.

In scoring the test, there is no guesswork about whether an answer is right or wrong. The rules and standards for correct answers are set up beforehand.

When a test is developed, given and scored carefully, and impartially, it is possible to say with some confidence that we have measured accurately, some part of the person.

Furthermore, if someone else administered the same test to another group of students at a later time and if the directions were followed precisely, any student could be compared with the scores of the group which was first tested, and his score evaluated. Therefore, to make a score on a standardized test meaningful, a lot of care must be used by teachers in giving the test.

Problem: Training 18,000 Teachers

The problem in Project Talent was to train over 18,000 teachers to give 30 standardized tests in the same way.

In some of the schools there were counselors with considerable training and experience in giving standardized tests. In these schools, the counselors could help the teachers give the tests. In other schools there were limited or no counseling facilities, and teachers had never given standardized tests before.

Thus, the amount of training necessary varied from school to school, depending on the experience of the staff and the facilities available. In some cases, the Regional Coordinators briefed the counselors, who then went over the details of testing with the teaching staff. In other cases, Regional Coordinators personally went over every detail of the two-day program to be sure that the procedures were completely understood.

The Tests Arrive

When test materials arrived at the school, they were essentially the responsibility of one person--generally the school principal or head counselor. It was his job to receive the materials and distribute them on the testing days; gather the materials up at the end of each half-day;

and organize the answer sheets for shipment to the Measurement Research Center, and of course, monitor the testing.

In all, 17 different instruments were developed for the schools' use in conducting the tests. These included not only the test booklets, and answer sheets, but teacher's guides, appendices to teacher's guides, aids to students in filling out the special answer sheets, student identification cards, special header sheets which were to be used for the return of answer sheets, school questionnaires, mailing labels, etc.

In addition, there was a manual specifically for the individual at the school who was responsible for the testing.

When the materials arrived at the school--and this was a truck load of boxes for the larger schools--they had to be sorted and counted for each homeroom. School people worked many extra hours just to make sure that the materials that were needed got to the right rooms and that every student got the same treatment as every other student in taking the tests.

All the effort of training the teachers, distributing and collecting the materials was for one purpose -- to insure that every student had the same opportunity to do well on the tests.

The Tests Are Given

All of the above preparations for giving the tests took place between December 1959 and March 1960, when the testing actually began.

Two methods of giving the tests were used. The most common method, of course, was for each homeroom teacher to

give the tests to his students. However, in quite a number of schools, the directions were read simultaneously to all students by one person through the public address system. In such cases, since there was only one timekeeper involved, there was no variation in timing within the school. Teachers, then, acted as proctors in each room. Although rare failures in the public address system did cause some difficulty, this method appeared to be very efficient.

The tests were given on either two full school days or on four half-days. When they were completed, the answer sheets were sorted, arranged, and shipped by the school. The next step--scoring and processing.

Giving and Taking the Tests--A Summing Up

Some Reasons for Participating. There is a feeling in many schools in this country that there is "too much testing". It is not uncommon, for example, for a half-dozen standardized tests to be scheduled for the school year. Teachers and administrators get the feeling that they are doing more testing than teaching.

In spite of this, when the Project was explained, these same people cooperated--almost unanimously. There are several reasons reported for this:

1. The potential values of the project were grasped readily by the school people. Counselors, especially, seemed to be sensitive to the total project, and realized its worth. They were very concerned that it should be done right, and were able to communicate

their concern to teachers and students. This was because many of the purposes of Project Talent are in support of guidance programs and problems. The values, which the test results will provide, in the form of guide posts for counseling to future generations of students and parents, are important to many professional people in the schools.

2. Many school people felt that participation was a patriotic contribution. Thus, many school people participated not so much because they felt they could get something out of the study, but because they felt they were cooperating in an important National Project.

3. Many schools have had little or no opportunity to do <u>any</u> testing, although they realize and are somewhat embarrassed about their lack of physical facilities, staff backgrounds, and personal experiences to do testing. These schools, of course, are generally small, and in isolated areas. The people in these schools are looking forward to any assistance that can be given to them from the outside.

A number of faculties voiced the sentiment that participating in the testing was a wonderful learning experience for them. They felt that the training given them in conducting so vast a testing program over such a short period of time was invaluable.

The Values for Students

High school students seem to have been enriched by this experience. Hundreds of students have written that they are looking forward with interest, amounting to eagerness, for the test results. Parents too, have written that they felt it was a rich experience for their children and would lead to serious considerations for future plans for their lives.

Most students are vitally interested in themselves, and were able to grasp the potential values of the Project. The experience of taking the tests widened their breadth of view, and, in many cases, made them aware of the need for planning for the present and future, and for attaining a more intelligent self-directive attitude.

Taking tests may help to make students more aware of one of the most important parts of understanding our fellow men--that people are different in many ways.

One guidance director reported that she firmly believed that the two days of testing in her school gave the students their best two educational days that they had had during the entire year. This comment was made despite the fact that it was not easy to arrange for the testing of 2000 students in two days; to be closely supervised in the process; and to have all this occur in an old building, outmoded by time.

The giving and taking of the tests required a tremendous amount of non-reimbursed labor on the faith that the Project will advance education, guidance, and human welfare in the United States. This would seem to be an immense vote of confidence in the future of our country, and in the men and women--our human resources--of the future.

THE ORIGINAL DOCUMENT WAS OF POOR QUALITY. BEST POSSIBLE REPRODUCTION FROM COPY FURNISHED ASTIA.

Chapter VI

Processing and Analyzing the Data

Since this is the age of automation--a major factor in making this study possible--the materials were planned to facilitate automatic data processing throughout. Answer sheets for tests and student questionnaires were set up in such a way that they could be scored, or otherwise processed, on the Iowa Electronic Test Scoring Machine or the Iowa Document Reader--two machines discussed later in this chapter which represent large-scale and completely up-to-date applications of modern technology to data processing.

After the basic data have been transferred in appropriate form to punch cards, primarily by these two machines, the data will be put on electromagnetic tape to permit subsequent processing on modern digital electronic computers.

Use of all these large scale technological marvels of machinery to process the test results and analyze the data might make it appear to the casual reader that there is danger of the individual getting lost in the huge mass of data--(almost a half-million students, two-and-a-half

million answer sheets, and five million IBM cards). But this will not happen. The reader may rest assured that modern data processing methodology makes it relatively easy to extract the data for specific students or small groups of students with specified characteristics from the mass of data, and to study these small groups of individuals. For instance if there turn out to be any Nobel Prize winners in the group twenty years from now, and surely there will not be many, it will be possible to locate their records and find out what they were like when they were in high school. Furthermore the use of high-speed computers makes it possible to incorporate into the data analysis a very large number of items of information about each individual--far more than one could even begin to consider if this kind of equipment were not available. The results therefore can provide a view in depth, not just a flat two-dimensional picture lacking in perspective.

Thus if astonishing progress had not been made in the past few years in the technology of processing and scoring the answer sheets by machine, Project Talent would have been impossible. Imagine obtaining about 100 scores per student for about 400,000 students by hand! Only slightly less fantastic, imagine obtaining that many scores on any scoring machine available before about 1958.

Automatic scoring of tests has come a long way since 1933, the year that IBM came out with the first electrical scoring machine. The operation of that machine relied on the fact that carbon conducts electricity. Since pencil "lead" is made of graphite, a form of carbon, it is possible to conduct electricity through a pencil mark on a test

answer sheet, thus closing an electrical circuit and causing a point to be added to the test score. The IBM machine based on this principle took the scoring of objective tests out of the horse-and-buggy age. But the jet age of scoring did not arrive until about 1958, when the Iowa Electronic Test Scoring Machine went into operation. This machine uses photoelectric cells (the basic components that cause a door with an "electric eye" to open automatically when someone approaches). In the case of the Iowa scoring machine, the answer sheet is scanned by an infra-red beam, and the photoelectric cell senses which of the several options for an item is the one which has been marked. The machine counts the number of "right" responses. Even if the student has ignored the instructions to mark only one answer per item, or has failed to erase an answer thoroughly when he changes his mind about the answer, the machine can cope with the situation. It does this by determining which one of the several responses that may be marked for an item is the darkest one; this is assumed to be the student's answer, the lighter marks presumably being incomplete erasures. If the two (or more) responses are equally dark, the machine treats the item as if it had been omitted. The machine scores at the rate of 6000 answer sheets an hour and can get 13 independent scores at a time, plus several composites, on each of these answer sheets. This means that 78,000 scores can conceivably be obtained in an hour not including composites; with the addition of whatever composite scores may be desirable the rate could conceivably go up to well over 100,000 scores per hour. Compare this with the 800 scores per hour that

would probably have been a top rate for machine scoring on any earlier machine. Then bear in mind that this amount of time would have had to be doubled since all the answer sheets would have had to be scored twice to insure satisfactory accuracy. (The Iowa Scoring Machine has its own built-in checks so that rescoring is not necessary.)

Now let's talk about a still newer development, the Iowa Document Reader. This is not another scoring machine; it has an entirely different purpose. Many of the items in the test battery are not combined into "scores". They are independent items of information. Consider for example the answers to the questions in the Student Information Blank. The individual answers must be preserved, not lost by being combined in some kind of composite with other items. Eventually, of course, selected sub-sets of these items might be combined in a designated way to answer certain very specific questions, but the items would still have to be kept separate so that other ways of combining them could also be tried if it should seem desirable. Then, of course, there is Record Form Z, which contains the student's name, address, date of birth, grade, and other information. Obviously all these separate items have to be recorded individually. The Iowa Document Reader takes care of the situation, by making it possible to transfer each individual response from the answer sheet or record form directly to punched cards. Like the Iowa Scoring Machine, it uses photoelectric cells to "read" the responses on the answer sheets--in other words to sense which response position is darkest for each question. Two items can be punched in each card column. Therefore, if ten of the card's 80 columns

are used for identifying information (the student's testing number, etc.), the student's responses to as many as 140 separate items can be punched simultaneously (two in each of 70 columns) on a single IBM card.

The initial processing of the Project Talent answer sheets took place at the Measurement Research Center, in Iowa City, since that is where the Iowa Scoring Machine and the Iowa Document Reader are located. The schools sent the marked answer sheets directly to Iowa City. Preliminary processing of the answer sheets, before they were fed into the machine, included checking whether they were fully identified as to grade, school, student number, and other necessary data, reconstructing answer sheets that had been badly damaged in transit (including one set that had apparently fallen into a bucket of apple green paint!) and other operations. Because these operations were all of a "clean-up" type, and also because this entire preliminary processing operation was being carried out in a building that not too long before had been a commercial laundry, the operation soon became known informally as the "Project Talent Laundry"; or in the interests of greater brevity, as "Operation LAUNDRY".

In addition to the activities mentioned above, "Operation LAUNDRY" had one other major function--namely the extraction of a four percent sample of all answer sheets except those returned too late for inclusion*. This four percent sample was to constitute a subsample on which tentative norms and other preliminary data analysis could be based. Student identification cards, numbered serially

---

* Answer sheets received at Measurement Research Center after 7 May 1960 were excluded.

from 100,001 to 670,000 had been printed. A block of these cards was sent to each school, for distribution (one card to each student). The six-digit number was the student's testing number, which he was to put on all his answer sheets. These testing numbers made it easy to get a representative four percent sample (or to be more precise, a sample that would be representative of those students whose answer sheets had been returned in time for inclusion). This was accomplished by selecting all answer sheets that had a testing number with 00, 25, 50, or 75 as the last two digits.

The two answer sheets that were put through the scoring machine yielded 26 scores apiece, or a total of 52 scores, plus 8 composites. Of these, 37 scores for each student were included in the reports to the schools. It was not considered appropriate to report the remaining scores, since all of them were either for tests of a highly experimental character or else they were of such a nature that they would have little utility for the schools. The 60 scores which were obtained on the scoring machine are listed in Appendix VI-A. The following scores are the 37 that were included in the reports to the schools.

1. Information - Vocabulary
2. Literature
3. Music
4. Social Studies
5. Mathematics
6. Physical Science
7. Biological Science
8. Aeronautics and Space
9. Electricity and Electronics
10. Mechanics
11. Home Economics
12. Information - Total
13. Memory For Sentences
14. Memory For Words
15. Disguised Words

16. English-Spelling
17. Capitalization
18. Punctuation
19. English Usage
20. Effective Expression
21. English - Total
22. Word Functions in Sentences
23. Reading Comprehension
24. Creativity
25. Mechanical Reasoning
26. Visualization in Two Dimensions
27. Visualization in Three Dimensions
28. Abstract Reasoning
29. Mathematics I-Arithmetic Reasoning
30. Mathematics II-Introductory
31. Mathematics I + II
32. Mathematics III-Advanced
33. Mathematics I + II + III
34. Arithmetic Computation
35. Table Reading
36. Clerical Checking
37. Object Inspection

For most of the tests the "raw score" obtained was a "rights score"-- i.e., the number of items answered correctly. For four of the tests, however, the "raw score" was obtained by a formula, in which the number of items attempted, as well as the number right, was involved. These four tests, and their formulas, are shown below.

Arithmetic Computation: $\text{Score} = R-3W = 4R - 3A$
Table Reading: $\text{Score} = R-W = 2R - A$
Clerical Checking: $\text{Score} = R-3W = 4R - 3A$
Object Inspection: $\text{Score} = R-W = 2R - A$

In the above formulas R represents number of items right, W represents number of items wrong, and A represents number of items attempted. (Thus $A = R + W$). The two forms in which each equation is presented (e.g., for Arithmetic Computation, "$\text{Score} = R - 3W$" and "$\text{Score} = 4R - 3A$") are mathematically equivalent. The explanation of these formulas will be discussed later.

The reports to the schools included not only the "raw scores" on the 37 tests but also the corresponding

"percentile"*. These were the preliminary norms that had been obtained on the basis of the four percent sample. For 33 of the 37 scores, percentiles had been obtained for each grade separately; in the case of the other four scores, which were in areas in which boys and girls customarily receive very different training, separate percentiles were obtained for each grade-and-sex combination. The four scores given this special treatment were:

Information Test: Aeronautics and Space
Information Test: Electricity and Electronics
Information Test: Mechanics
Information Test: Home Economics

Tables showing the percentiles corresponding to raw scores on each of the 37 measures appear in Appendix VI-B.

Additional preliminary analysis of the four percent sample included the determination of the intercorrelations among 58 variables for boys and girls separately, and separately by grade. This meant a total of 8 tables of intercorrelations. These tables are presented in Appendix C.

In addition to the 60 scores that were obtained on the scoring machine for everybody, 46 additional scores were obtained for the four percent sample, primarily for research use. These additional 46 scores are indicated in Appendix VI-A.

Most of these additional raw scores, the reader will observe, consist of number of items attempted on various

---

* A percentile is a numberical value which indicates how someone stands on a test or other variable in comparison with a specified group. The 80th percentile, to take a specific example, is the raw score point below which 80% of the individuals in the basic group score and above which the remaining 20% score. Thus a percentile can have any value between 0 and 100.

tests. The purpose of determining these counts was to see what the effect would be of using a raw score which had been obtained by taking into account both the number of items answered correctly and the number answered incorrectly, instead of just using the number answered correctly, without making any differentiation between omitted items and items answered incorrectly.

When raw scores obtained by formula from number right and number wrong are used, instead of just the number right, it is usually for one of two reasons. The more common of these reasons is that it seems desirable to use raw scores which have been "corrected for chance". Somewhat less frequently, the content or the purpose of a test is of such a nature that more useful results can be obtained by using a scoring formula that differs somewhat from the standard correction-for-chance formula.

The so-called "correction for chance" is a procedure for taking into account not only the number of items a student answers correctly but also the number he answers incorrectly (as opposed to the ones he omits). The basic assumption underlying the correction formula is that if a student gets an answer wrong it is quite likely that he guessed; and that if he did not engage in out-and-out guessing, he obtained the wrong answer on some other basis which would occasionally net him a right answer for a wrong reason. For instance, in the case of five-choice items it is probably that for every four items that a student answers incorrectly he answers one item correctly by chance alone, or possibly not quite by chance but nevertheless for a wrong reason. In other words it is assumed that for every

four items a student answers incorrectly he answers one correctly without really <u>knowing</u> the answer. If this assumption is correct, it is entirely equitable to "correct" a student's score on a test with five-choice items by subtracting from the number he answers correctly one-fourth of the number he answers incorrectly, since this presumably deducts the number of items he answered correctly without really knowing the answer. The chief reason for scoring the four percent sample two ways ("corrected" scores and number right) was to permit a comparison of the results obtained with the two procedures*.

As has been indicated above, when a scoring formula other than "Number Right" is used, it is not always the correction-for-chance formula. For instance in the case of the four tests on which the scores are routinely being obtained by formula for everyone (the Arithmetic Computation, Table Reading, Clerical Checking, and Object Inspection tests), the formulas used do not yield a simple "correction for chance". If that type of score were to be obtained, the formula would be

$$\text{Score} = R - \frac{W}{4}$$

for the tests with five-choice items (Arithmetic Computation, Table Reading, and Object Inspection) and

$$\text{Score} = R - W$$

for the test with two-choice items (Clerical Checking). In the case of these four tests, wrong answers are much more

---

* At the time this book goes to press, this comparison has not yet been completed. The basic data for this comparison will include the correlation of the "Rights score" on each test with the corresponding "corrected score", and also tables of intercorrelations among corrected scores, to be compared with intercorrelations among Rights scores. (These latter intercorrelations appear in Appendix C.)

likely to be due to carelessness than to anything even remotely akin to guessing or "chance". Theoretical considerations and past research indicate that in tests of this sort, where speed and accuracy in performing a somewhat routine task (such as a routine clerical task) are being tested, more useful results are usually obtained if a stiff penalty is imposed for inaccuracy than if the deduction for errors is merely sufficient to counteract items answered correctly by chance. The scoring formulas that are being used for the four tests concerned are ones that on the basis of past experience seem likely to give good results. However the choice of these formulas involves no irreversible commitment. The magnetic tapes (described later in this chapter) which will contain the permanent record of scores will preserve the "Number Rights" score and the "Number Attempted" score for each student, so that if future data analysis suggests that better prediction might be obtained with different scoring formulas from the ones chosen, the basic data will be available for applying the preferred formula. In effect this amounts to treating "Number Right" and "Number Attempted " as if they were scores on two separate tests, and giving each of them whatever weight (positive or negative) gives the best prediction of whatever one happens to be trying to predict.

As has been implied above, what the best scoring formula is for a test depends on the purpose for which the test results are to be used. A word or two on how this works out seems desirable. A student's failure to answer a test item correctly is probably due to one of the following reasons:

1. Lack of knowledge of the correct answer. (This is not a highly likely reason in the case of the four tests for which a scoring formula other than the correction-for-chance formula is being used.)
2. Carelessness (This will probably result in wrong answers, rather than omitted items.)
3. Not having enough time to reach the items. (This, of course, results in omitted items, not wrong answers.)

Since the first of these three reasons is not particularly applicable to the four tests we are concerned with at the moment, the ones for which a scoring formula other than the correction-for-chance formula is being applied routinely, let us limit further discussion to the second and third reasons.

For prediction of success in jobs calling for careful performance of a somewhat routine task where carefulness and accuracy may be considered more important than speed, a formula which has the effect of penalizing carelessness (as manifested by wrong answers) more severely than slowness (as manifested by unreached items) seems likely to prove appropriate. This means a formula which imposes an extra-heavy penalty for wrong answers. It is of course conceivable, however, that for other types of jobs, in which speed may be far more important than accuracy, different scoring formulas which imposes a heavier penalty for failure to reach an item more heavily than for a wrong answer would give more accurate predictions. If such is the case, we shall find out about it in the follow-up.

Perhaps some of our readers are becoming a little concerned by now because so much has been said about kinds

of scoring and data analysis and so little about precautions to make certain that the data have not been rendered meaningless by administrative and other errors, of one sort or another. In a project as huge as this one, it is inevitable that not everyone who participates in the task of administering the tests will follow directions perfectly, and that there will be some schools in which things do not go smoothly. Fire alarms can occur at unexpected times; clocks can break down; so can stopwatches; test administrators can turn to the wrong page in the Teachers' Guide for use in administering the Project Talent Test Battery and therefore use the wrong time limits; other things too can conceivably go wrong. How, then, can one be sure that the data will not be hopelessly distorted by resultant inaccuracies, and that conclusions consequently will not be hopelessly wrong? The answer is that great pains have been taken to get as complete and accurate a record as possible of all irregularities that occurred in the testing, so that if these irregularities, in the considered judgment of the Project staff, are sufficient to invalidate the results for the students affected, the scores involved can be labeled invalid, and eliminated from the data analysis. The Test Record Form, which is shown in Figure 1, is the primary source of the necessary information for this purpose. Not only does it provide a record of all irregularities that occurred in the testing, but it also indicates the testing numbers of the students affected.

A supplementary form, the "Local Coordinator's Record of Identification Numbers", which is shown in Appendix D, serves several purposes, one of which is to supplement the

Test Record Form in indicating which students are in the group affected by a particular irregularity.

However the data processing procedure was such that it was not feasible to incorporate the information from the Test Record Forms as early in the chain of operations as would be necessary if this information were to be used in the analysis of the four percent sample. Therefore in that analysis other means had to be used for screening out questionable scores. The interim screening mechanisms used were documents called Group Identification Sheets. There were two kinds--Type I, to be used with Answer Sheets A and C (the scoring machine answer sheets), and Type II, to be used with Answer Sheets B1 and B2 and Record Form Z (the document reader sheets). They are shown in Figures VI-2a and VI-2b, respectively.

The schools had received instructions to separate those answer sheets that contained tests administered under irregular conditions behind appropriately identified Group Identification Sheets, one for each grade group. The corresponding packs of answer sheets containing only tests that were free of irregularities in administration were to be put behind other appropriately labeled Group Identification Sheets.

All answer sheets in the four percent sample which had been identified as containing irregularities according to the Group Identification Sheets were eliminated before the results were analyzed and the preliminary norms obtained. Thus the entire answer sheet was eliminated even if only one test on it involved irregularities. (This very conservative policy was necessitated by the fact that the

Figure VI-1

TEST RECORD FORM - PROJECT TALENT

One of these forms is to be filled out by the teacher or other examiner for each room in which tests are administered.

School Code__________ School____________________

City________________ State________________

Teacher or examiner:________________________

Tests Timed with (check one):

_____a. Stopwatch

_____b. Interval Timer

_____c. Clock with sweep second hand

_____d. Watch with sweep second hand

_____e. Other (specify):__________

Did any irregularities occur in testing ?

No________

Yes________

If so, to whom did they apply ?

☐ Entire group

☐ Only to individual students, (who are identified on the back of this sheet by their 6-digit testing numbers).

Describe all these irregularities on back of this sheet in detail.

✓	Grade	No. of Students
	9	
	10	
	11	
	12	
	Other	
Total		

Identification Cards	No. of cds.	Testing Numbers (6 digit nos.) Lowest	Highest
Used			
Unused			
Total supply			

Test Session	Date Mo.Day Yr.	Check One AM	PM	No. of Students Pres.	Abs.	Counts chkd. by (Sign initials) Booklet		Ans.Sht.		Ident. Info. Chkd.
1						A		A		
								Z		
2						B		B1		
								B2		
3						C1-X		C		
						C1				
4						C2				

Bklt	Ans. Sheet	Test *			No. of Min	Time Start	Time Stop	Act. Time	Do not mark these columns			
										+	-	
Session 1												
A	A	Information - I			90				1			
		BREAK			3	/////	/////	/////	/////	/////	/////	/////
		Student Act. Inv.			20				2			
		Preferences			3				3			
-	Z	Master Record Form			/////	/////	/////	/////	4			
Session 2												
B	B1	Interest Inventory			20				5			
		Information - II			35				6			
	B1,B2	BREAK			3	/////	/////	/////	/////	/////	/////	/////
		Student Info. Blank			80				7			
	B2	Paragraphs	#395		5				8			
	B1		#396		5				9			
Session 3												
C1-X	C	Mem. for Sent.-Study			6				10			
		Arith. Computation			9				11			
		Memory for Words	Study		2				12			
			Practice		2				13			
C1			Test		4				14			
		Mem. for Sent.-Test			10				15			
		Math	I.Arith.Reas.		12				16			
			II.Introductory		24				17			
			III.Advanced		14				18			
		BREAK			3	/////	/////	/////	/////	/////	/////	/////
		Eng.	I.Eng.Usage		10				19			
			II.Eff.Exp.		10				20			
			III. Punct.	a.Punc.Mrk	9				21			
				b.Sent.Str.	5				22			
			IV.Spelling		12				23			
			V.Capitaliz.		6				24			
Session 4												
C2	C	Abstract Reasoning			11				25			
		Mechanical Reasoning			11				26			
		Disguised Words	Directions		1				27			
			Test		3				28			
		Creativity	Dir.		1				29			
			Test		20				30			
		Clerical Checking			3				31			
		Vis. in 2 Dimen.			4				32			
		BREAK			3	/////	/////	/////	/////	/////	/////	/////
		Reading Comprehen.			30				33			
		Vis. in 3 Dimen.			9				34			
		Word Funct. in Sentences	Dir.		2				35			
			Test		15				36			
		Table Reading			3				37			
		Object Inspection			3				38			

I certify that the above record is accurate, to the best of my knowledge.

Signed:________________________
Teacher or examiner

* Times given are for actual tests (not for directions, etc.) unless otherwise specified.

© Copyright 1960. All rights reserved
University of Pittsburgh
Pittsburgh, Pennsylvania

Printed in U.S.A.

TALENT Testing Program
GROUP IDENTIFICATION SHEET - Type I
For use with Answer Sheets A and C

Directions for filling out this form appear in the Local Coordinators' Manual in Paragraph IX. Please read these directions and follow them carefully.

Please print:

1. Name of school ______________________ School Code* ________

City ______________________ State ________

2. Group which this sheet accompanies (check one):

___ Grade 7

___ Grade 8

___ Grade 9

___ Grade 10

___ Grade 11

___ Grade 12

___ Grade 13

___ Special Sample (15 year-olds)

___ Other (specify) ______________________

3. Answer sheet which this sheet accompanies (check one):

___ A

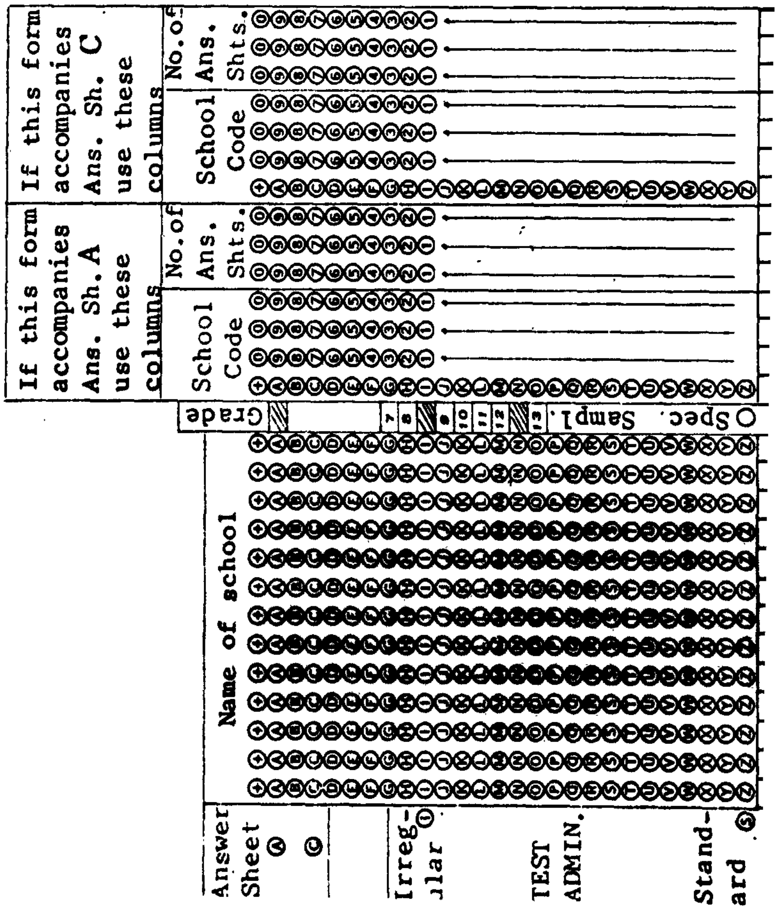
___ C

4. The test administration for these cases was:

___ Standard (i.e. it involved no irregularities)

___ Irregular (i.e. it involved irregularities reported on the Test Record Form)

5. Number of answer sheets in this pack ________



* School code is always a letter followed by three digits.

PLEASE -- Keep this form clean and free from stray marks. Do not fold it or bend the corners.

n Figure VI-2b

TALENT Testing Program
GROUP IDENTIFICATION SHEET - Type II
For use with Record Form Z and Answer Sheets B1 and B2

Directions for filling out this form appear in the Local Coordinators' Manual in Paragraph X. Please read these directions and follow them carefully.

Please print:

1. Name of school ______________________ School Code* ______________

City ______________________ State ______________

2. Group which this sheet accompanies (check one):

___ Grade 7

___ Grade 8

___ Grade 9

___ Grade 10

___ Grade 11

___ Grade 12

___ Grade 13

___ Special Sample (15 year-olds)

___ Other (specify) ______________________

3. Answer sheet which this sheet accompanies (check one):

___ Z

___ B1

___ B2

4. The test administration for these cases was:

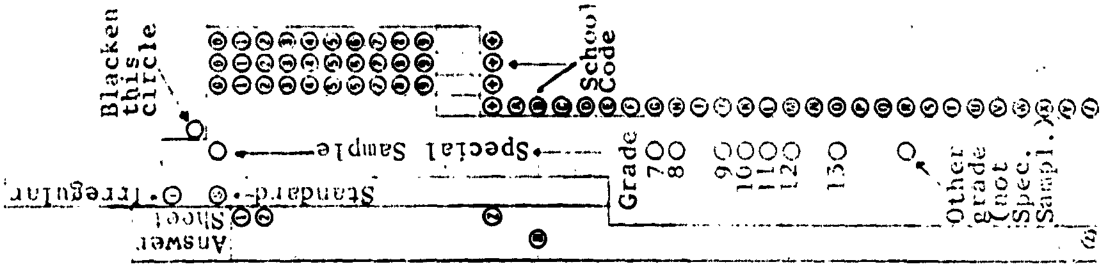
___ Standard (i.e. it involved no irregularities)

___ Irregular (i.e. it involved irregularities reported on the Test Record Form)

5. Number of answer sheets (or Record Forms) in this pack: ______________

* School Code is always a letter followed by three digits.

MAKE NO STRAY MARKS ON THIS SHEET.
PLEASE -- Do not fold this sheet or bend the corners.



Group Identification Sheet does not contain the details about the nature of the irregularities. After the Test Record Forms have been fully processed and the data they contain incorporated with test results, it will be possible to salvage data on "irregular" answer sheets by eliminating only those scores that are affected by the irregularity.)

After the basic data have been transferred in appropriate form to punch cards, primarily by the Scoring Machine and the Document Reader, the data will be put on electromagnetic tape (often called magnetic tape) to store it compactly and to permit subsequent processing on modern digital electronic computers.

Most of the data processing will be done on the giant computer at the University of Pittsburgh. This machine is an "IBM-7070", and even in comparison with most other electronic computers currently in use it is capable of prodigies, in terms of volume and speed of data handling and also in terms of versatility. The 7070 is still a comparatively new machine--comparatively new even in a field where technological progress is so rapid that yesterday's eighth wonder of the world is today's commonplace and tomorrow's obsolescent relic.

The 7070 is a solid-state (transistorized) fully "buffered" high-speed general-purpose decimal computer that can "remember" as many as 99900 digits. The memory is in the form of "core storage". The computer uses high-density tape and it can run as many as six reels at a time. In addition to the tape input and output it also has card input (a high-speed card reader), card output, and a printer. Those of our readers who happen to be "computer buffs" will

have no trouble understanding the technical terms in the preceding sentence. As for the rest of our readers we apologize for not explaining these technical terms in detail but limitations of space, and perhaps of reader patience, seems to make such a digression inadvisable. We will therefore pause only long enough to mention one additional feature, which is one of the very newest wrinkles in computer design--so new, in fact, that the 7070 is one of the very few computers beyond the drawing-board stage that have it. This is the fact that the computer's "automatic priority processing" feature makes it capable of running several completely independent programs at the same time. This feat would perhaps be somewhat analogous to that of a person who could compose a sonnet in French and commit it to paper with his right hand, while using his left hand to solve a mathematical equation, and at the same time could carry on a sensible conversation about the latest best-seller. Though this hypothetical stunt is quite a bizarre one, and not notably practical, the 7070's corresponding capability for unrelated simultaneous activities can be put to really practical use. It was fortunate that this machine became available just in time for Project Talent.

Now let's get back to the magnetic tape on which scores, item responses, and other data for each of the Project Talent students have been recorded. Some of our readers may be surprised to learn that the record is an invisible one! Unlike the information punched in cards, which is represented by very visible rectangular holes, the information recorded on magnetic tape is represented by invisible patterns of magnetized spots. Since there is no

need for the record to be visible to the human eye, it can be extremely compact. Thus a 2400-foot reel of half-inch-wide tape has a diameter of only ten and one-half inches but can contain about 15 million individual characters. (Each character is either a letter of the alphabet, one of the ten digits, or a special symbol.) Now to give the reader some idea of the speed with which the 7070 can handle these tapes perhaps it will suffice to mention that an entire reel, containing about 15,000,000 characters, can be run through the machine in well under ten minutes.

Compactness of storage and speed of processing have been mentioned as advantages of magnetic tape. A third major advantage is flexibility. It will be necessary to bring the tape up to date after the one-year, five-year, ten-year, and 20-year follow-ups. This can be handled automatically if the machine is "programmed" appropriately. (i.e., if the machine is given appropriate instructions appropriately expressed in machine language).

Most of the rest of this chapter will be devoted to plans for data analysis. Before we concern ourselves with methodology a brief discussion in broad general terms of the kinds of analysis contemplated seems in order.

The kinds of analysis that are possible are of course implicit in the instruments developed and the kinds of data collected. It will be possible to find the answers to many questions about what the typical high school student is like--his aptitudes, abilities, activities, interests, background, and plans; also what the atypical high school student is like. The sample is sufficiently large that it will be possible to identify reasonable numbers of students with highly superior abilities, students with plans to go into occupations for which few people can qualify (e.g., nuclear physicist), students with unusual combinations of traits, unusual interests, etc., find out what they are like, and see what happens to them. In other words extremes as well as means can be studied.

But more than that, the Project is not limited to studying secondary school students. Because an entire age group--all 15-year-olds whether in secondary school or not--was

sampled, it will be possible to draw some inferences in regard to this much broader and more general base--which in a sense is a random sampling of "people-in-general", since everyone who survives past his fifteenth birthday is a fifteen-year-old at one time or another, while not everyone is at some time a secondary school student.

Furthermore, the Project will not be limited to studying people (either "people-in-general when they are fifteen years old" or "people-who-go-to-secondary-school"). In addition special studies will be done of schools. With 1357 schools being tested and with procedures for weighting the schools so as to reconstruct the "population" of secondary schools in the United States as a whole, it will be possible to find out what typical schools are like, how schools differ, and how the characteristics in which they differ are related to their success in educating young America. Perhaps most important of all, the results may even point the way to how some schools can improve their program--what ways of utilizing available resources (monetary and other) will result in the greatest increment in the degree to which they succeed in their objectives.

Almost all of the data analysis still lies ahead. Since this is so, there seems to be little point in discussing methodology in great detail at this stage. However, general procedures and problems that are expected to apply throughout may be discussed with profit.

One of the general problems is how the total group of a half-million cases shall be divided into sub-groups for analysis. Obviously it is undesirable to throw a half-million students together in an undifferentiated mass. Therefore

it is expected that in general whatever analyses are done will be done separately by grade, and in most cases the boys and the girls within a grade will also be analyzed separately. Where these breakdowns still give larger groups than it seems desirable to handle as a unit, the group can readily be broken down into ten smaller groups, by use of the ten subsamples of schools which the reader will recall were described in Chapter 3 (on page_____). In addition to the breakdowns described above, the group can also be divided on a geographic basis, so that regional norms can be obtained; norms will be obtained for all sorts of special groups as well, to supplement the national norms.

In the case of the special study of the 15-year-olds, the boys and girls included will come from two sources--students in the regular sample who were in grade 9, 10, 11, or 12 in one of the randomly selected high schools (or an associated junior high school), and the special group of 15-year-olds who were not in high school at all. Appropriate weights will have to be applied to the data for students from these two sources before they can be combined. This is necessary to take account of the fact that the group not in high school represents only one-tenth as many school districts as the group in high school.

Another problem which has implications for many aspects of the project is that of follow-up. Students will be followed up one, five, ten, and twenty years after graduation from high school. The authors feel quite optimistic about the prospects of getting excellent cooperation from the students. However it must be recognized that over the course of the years the follow-up will become

progressively more difficult, as more members of the group change their addresses and forget to notify the Project Talent Central Office, or change their names, or for some other reason become difficult to locate, or disinclined to report their current activities. One device which the authors have great hopes will recover some of these lost cases is to reverse the follow-up procedure. The reverse procedure, which is here designated the "follow-back procedure", involves finding out the names and approximate ages of the members of any specified group which is of particular interest ten or 20 years from now, finding out which members of the group are of the right age to have been tested in Project Talent, and then acting on the assumption that approximately 5% of these people were in Project Talent.(This is not an unreasonable assumption, since approximately 5% of today's high school students are included in the Project.) Let's see how this follow-back procedure would work in a specific instance. Suppose that about twenty years from now there is an interest in studying persons convicted of certain kinds of crimes, and finding out what their characteristics were when they were in high school. Lists of persons convicted of the crimes could be inspected, and the names of all such persons who are in the right age range could be extracted. IBM cards could be punched for these persons, showing their name, date of birth, sex, and other appropriate identifying information. The computer would then be set up to "search the file" for these names. Presumably about 5% of them would be located. The SAI scores, responses on the Student Information Blank and Interest Inventory, test scores, and other relevant data

for these individuals could then be studied intensively, and compared with the corresponding data for the total group in the study or for any segment of it. This would provide information on what these people were like when they were in high school, what backgrounds they had, and in what respects, if any, they differed from their contemporaries.

In previous parts of this book there have been references to the problem of assigning appropriate weights to the data to adjust for the fact that different sampling ratios were employed in different kinds of schools and also to adjust for various other circumstances, such as the fact that a number of schools that were selected in the sample declined to participate in the project. To allow for foreseeable contingencies, three different sets of school weights, designated Set A, Set B, and Set C, respectively, have been determined. Eash is appropriate in a different set of circumstances. Weight A is the weight that is applied to the data when the characteristics of students are being studied. Use of these weights makes it possible to reconstruct from the sample an approximation of the total population of high school students--or any segment of this population that is of particular interest. Weight B, which is exactly the same as weight A for most schools, is applied when the characteristics of schools are being studied, and it is desired to infer something about the total group of American high schools on the basis of the schools that participated in the project. (Weight B differs from weight A only in the case of those schools where only a fraction of the students were to be tested.

Page ____ of ____

LOCAL COORDINATOR'S RECORD OF IDENTIFICATION CARDS

School ______________________ City ______________________ State ______________

Identification Cards Received: From __________ to __________ Test Date __________ __________ __________
Lowest No. Highest No. Month Days Year

Grade	Room	Other Group Desig- nation*	Teacher Administering the Tests	No. of Students to be Tested	Cards Issued to Teacher			Cards Issued to Students			Unused Cards		
					Quantity	Ident. Card #'s		Quantity	Ident. Card #'s		Quantity	Ident. Card #'s	
						From	To		From	To		From	To

* Use of this column is optional

Signed ______________________
Local Coordinator

The only schools in this category were the public schools in New York City and those in Chicago.) Weight C, like Weight B, is designed for use where it is desired to draw inferences about the total group of American high schools. However, unlike weight B, weight C is used when the data being studied are available for all schools that were selected in the sample, rather than just for those schools which agreed to participate. It should be pointed out here that even the schools that declined to participate were in most cases willing to fill out the three school questionnaires. Therefore, when the data from these questionnaires are being used and test scores are not, weight C may be the appropriate one. [redacted]

CHAPTER VII

BACKGROUND AND DESCRIPTION OF THE TESTS

In Section I of this book, the rationale for the project as a whole, its raison d'etre, was discussed. Section II contains the rationale for the procedure used. In Section III the rationale for each of the tests and other instruments will be presented.

In this chapter, each test is discussed in turn, in terms of its purpose, its historical antecedents, and its characteristics. The section on historical antecedents includes not only a description of earlier tests designed to serve the same general purpose but also a description of ways in which these tests have been used successfully in the past--the kinds of jobs for which the abilities they measure have proved relevant.

Testing has a long history--as long as the history of schools and perhaps even longer. But to keep this discussion of historical antecedents of the Talent tests within reasonable bounds we shall limit it in general to standardized objective tests.

The discussion of the Project Talent version of each kind of test includes a brief description of the item type, sample items, and a discussion of whether the items used are essentially parallel to those in other tests having the same general purpose, and if not, to what extent they differ. The reader may wish to refer back to Table IV-2 for a summary of certain features of the tests including number of items and number of options for each. All items in all tests are of the multiple-choice type, and the reader will note from this table that almost all of them are five-choice items.

A. Information Test

Purpose of the Test

The Information Test is really several tests rolled in one package and printed in the same booklet. Devotees of "Alice in Wonderland" and "Through the Looking Glass" will recall Humpty Dumpty's use of "portmanteau words"--words that have several seemingly disparate meanings at the same time. In much the same way, the Information Test has what might be called a "portmanteau purpose". The test may be scored to contribute to the measurement of breadth of the student's general information, his vocabulary level, and the amount of information he has in many specific areas (including some areas in which amount of information has been found to be an excellent index of aptitude, and other areas, of an academic nature, in which amount of information gives some indication of level of scholastic achievement); the test also helps identify mentally retarded students, and others who have a serious reading disability--those who read badly enough to be considered

functionally illiterate; and students whose attitude toward the test-taking task is flippant, or apathetic, or otherwise leaves much to be desired; and students whose scores may be pulled down by careless errors; and finally it helps determine whether the student's general attitude in the face of unexplained phenomena leads him to take a scientific view in evaluating hypotheses, or whether his general attitude makes him receptive to explanations which are superstitious or otherwise illogical.

This is quite a bit to fit into a single test--even if it is a long one, as the Information Test certainly is. The one thing that makes it possible to achieve all these ends is the fact that most of the items contribute to more than one of the purposes. For instance every item, regardless of what else it measures, also contributes to the total "general information" score. The items that serve to identify illiterates also help to identify the inadequately motivated students--provided that the entire pattern of scores is taken into account and that the results are properly interpreted.

Historical Antecedents

Because the purposes of the Information Test are so varied, and because the authors know of no previous test designed to serve all these purposes, the historical antecedents will be discussed for each purpose separately.

For purposes of this discussion we shall define information test items quite broadly, to include items testing almost any type of information: information in academic areas, information in non-academic areas, information about hobbies and recreational activities; information acquired

through formal instruction, information acquired in daily living, information acquired through voluntary reading; information about people, places, events, things, activities, word meanings. Items for the last-named category, information about word meanings, are usually called "vocabulary items"; but the distinction between items specifically calling for knowledge of word meanings and items calling for other kinds of information is rather an arbitrary one, and we shall not trouble ourselves about it at this point. (We shall resume the discussion, however, in a later section, when we discuss "vocabulary scores".)

Our definition of information test items is sufficiently broad to include not only items which are purely verbal in character but also items which are largely pictorial. In the latter category for instance, would be an item asking:

"Which of the following animals is a jaguar?"

and for options showing drawings of five kinds of wild animals.

1. <u>Measures of breadth of general information</u>. Information tests covering a wide range of topics have been used for a long time as components of general intelligence measures. They were used for this purpose as far back as World War I, when the Army Alpha Examination was administered to all inductees into the Army who could read and write. This was the first mass program for mental testing that ever took place. The Army Alpha was a standardized objective test designed for group administration. One of its eight subtests was a test of information, which covered a very wide range of subject matter. To give an idea of its breadth of coverage, it included items testing information

on animals and birds, physical science, biological science, farming, mechanics, social studies, electricity, medicine, games, business, foods, airplanes, sports, guns, law, literature, art, and other topics. Arthur S. Otis, who played a leading role in the development of the Army Alpha, later used information items as one of the chief components of his widely used intelligence tests (the Otis Self-Administering Tests of Mental Ability and the Otis Quick-Scoring Mental Ability Tests).

Use of information tests as an indicator of general intelligence level was thus one of the earliest uses of such tests and it is still a very important one. In this type of use only the total score is of concern; subscores showing amounts of information in specific areas are irrelevant and are usually not obtained. A wealth of empirical data has accumulated concerning the use of general information as a measure of intelligence, and predictor of college success.

2. Measures of interest. In the 1930's there was some experimentation with wide-range information and vocabulary tests as a way of measuring relative interests. For instance there was the Michigan Vocabulary Profile Test ( ),' developed by E. B. Greene in 1937. This test consisted of eight subtests, measuring vocabulary in the following eight areas: Human Relations, Commerce, Government, Physical Science, Biological Science, Mathematics, Fine Arts, and Sports. The total score on this test clearly provides a measure of general vocabulary, but the innovation which distinguished this test from its predecessors was the recommendation made in its manual that the profile of subtest scores,

in other words the pattern of these eight scores, be used as a clue to the person's relative interests in different areas and as a partial indication of his relative aptitude for work in these areas.

The use of differential levels of information in different areas as a measure of relative interests is based on the premise that the individual will tend to know most in areas where his interest is greatest. Thus if the profile of his subscores shows peaks and valleys, these extremes may be taken to suggest areas where his interests, and thus perhaps his ability to succeed, are greatest and least. (Implicit in all this, of course, is the assumption that the subtest scores have all been converted to the same scale, to make them directly comparable, since raw scores presumably would not be.)

The use of information tests to suggest whether an individual has sufficient interest in a field to succeed in it was greatly expanded and systematized during World War II in the Army Air Force Aviation Psychology Program, under the direction of John C. Flanagan. It proved very successful in the battery of Aircrew Classification Tests designed to predict success as an Air Force pilot, bombardier, or navigator. This battery initially included a Technical Vocabulary Test, which provided separate scores for the three positions named. The test was later expanded to include other types of general information.

A battery later developed by the American Institute for Research for use in selecting airline pilots and flight engineers made use of many principles and methods which had proved useful in the Air Force, among them the use of

information tests to provide an "operational" measure of interest in flying--a measure based not on mere verbal expression of interest but on interest that has been demonstrated, by the acquisition of information. The extent to which information about flying and other kinds of information deemed relevant indicate a real interest rather than merely the acquisition of a wide range of information through the operation of general intelligence can be determined by comparing a score based only on relevant information with a score based on other information, or with the total information score.

3. Measures of information in specific areas. In addition to their usefulness in indicating relative interest in different areas, subscores on a general information test are also useful per se, as direct measures of amount of information acquired in certain specialized areas--for instance physical science, fine arts, sports. The Michigan Vocabulary Profile Test, which has already been mentioned in connection with the measurement of interest, provides such scores. The distinction between academic areas and non-academic areas is a rather fruitless one in the case of general information tests, since the information tested often includes types of facts that are just as likely to have been learned out of school as from formal instruction.

A second distinction that is by no means as clear as it is sometimes thought to be is the distinction between information as a measure of "aptitude" and information as a measure of "achievement". Many tests may be regarded as having elements of both. The fact that a person has acquired a lot of information about a specialized area

indicates a kind of achievement. It also indicates that he has the aptitude, and insofar as the information was acquired through voluntary activity, it indicates interest in the area. In the sense that past performance predicts future performance, achievement in a field is an index of aptitude for it and thus information in an area may indicate both achievement and aptitude. Nevertheless there are certain areas in which these joint achievement-aptitude tests that measure amount of information have proved outstandingly useful as predicters or indicators of job aptitudes. Among these areas are: (a) mechanical information, (b) electrical and electronic information, and (c) aeronautical information. Each of these will be discussed separately in a later section of this chapter since the Information Test items in these three areas may be regarded not only as contributing to the total Information score but also as constituting subtests in three areas of specialized information which are embedded in the Information Test.

We have already suggested that the distinction between "vocabulary items" (items testing knowledge of technical vocabulary) and "information items" (items testing possession of other kinds of information, in the same field) as not an important one. Actually the two go hand-in-hand. Students who do well on either of the two types of items (technical vocabulary items and short concise items testing specialized information) tend to do well on the other, and students who do poorly on one tend to do poorly on both.

4. **Measures of vocabulary.** The reader should not infer from the foregoing discussion that the authors regard scores of general vocabulary level as unimportant. Quite the contrary! The size of a person's general vocabulary has long been recognized as one of the best indices of his level of general intelligence--or at least of that very important aspect of general intelligence that is called "verbal intelligence". Vocabulary tests of one sort or another have played a prominent role in intelligence tests ever since the first intelligence test, the Binet scale, was developed back in 1905. Almost every intelligence test except those specifically designed to be non-verbal contains either a vocabulary subtest or vocabulary items intermingled among other kinds.

Just as there are more ways to kill a cat than by choking it with butter, there are more ways to measure a person's vocabulary than by putting him in a conversational group and seeing how many people he can out-talk or by counting the number of people over whose heads he can talk! Standardized paper-and-pencil tests are a much better way. And even these can take many forms. Verbal analogy items, items calling for words opposite in meaning to given words, items calling for synonyms, and items calling for matching a word with a definition have all been used. Samples of some popular types of vocabulary items are shown below.

a. Verbal analogy items

Item 1. "Place" is to "time" as "point" is to

A. hour
B. minute
C. moment
D. eternity
E. now (The answer is C.)

Item 2. Hand: gauntlet: sword: __________?

A. attack
B. armor
C. shield
D. spear
E. sheath

(The answer is E.)

b. Opposites items (antonym items)

Directions: Choose the word that is most nearly opposite in meaning to the given word.

Item 3. Hope

A. Dislike
B. Despair
C. Wish
D. Fail
E. Criticize

(The answer is B.)

c. Synonym items

Directions: Choose the word that has most nearly the same meaning as the given word.

Item 4. Apparel

A. Appearance
B. Illusion
C. Luggage
D. Clothing
E. Elegance

(The answer is D.)

Item 5. Mystify

A. Cheat
B. Puzzle
C. Vaporize
D. Scent
E. Enlarge

(The answer is B.)

d. Same-opposite items

Directions: Each item consists of four words. Find the two that are most nearly the same in meaning or opposite in meaning.

Item 6

A. Adjust
B. Admit
C. Adapt
D. Adept

Item 7

A. Exotic
B. Superficial
C. Ingrained
D. Excellent

(The answer to Item 6 is A and C, and to Item 7 is B and C.)

e. Definition items

Item 8. Suffrage is the right to

A. petition.
B. vote.
C. assemble.
D. worship.
E. work.

(The answer is B.)

These varied item types all give quite similar results. It is true that the verbal analogy type of item (and to a lesser extent the same-opposite type) may involve a little more reasoning ability than some of the other types, but even so, what any of them measures is mostly vocabulary.

5. Identification of illiterates and persons with severe reading disabilities. Among the items in the Information subtest of the Army Alpha were some extremely easy ones, that had been put in for the purpose of screening illiterates. For instance there was an item testing knowledge of the fact that it is more likely to snow in the winter than in the summer; and an item testing knowledge of the fact that eggs come from hens. Other tests, too, have contained very easy items to insure that anyone who could read and write could get better than a zero score.

The usual procedure has been to incorporate these extra-easy items with the items of normal difficulty in a single score. It has not been customary, however, to obtain a separate score on the extra-easy items, which is the procedure that was followed on Project Talent, in an effort to identify students whose reading skills are so inadequate that their scores on the rest of the Information Test and on other tests in the battery would be invalidated by failure to comprehend the items.

6. <u>Measures of motivational factors</u>. A poor score on items that are so easy that anyone who can read and write should get a perfect score usually indicates either illiteracy or else a severe reading disability characterized by abnormally slow reading or by grossly inadequate comprehension of what is read. If these explanations of the poor score on extra-easy items do not fit, the next more probable alternative explanation is that the attitude of the person taking the test leaves much to be desired. In this event the low score may be due to flippancy (manifested by deliberately marking the wrong answers) or to apathy and indifference (resulting in lackadaisical behavior in marking the answers).

Thus the same items that are used to spot illuterates and "non-readers" may also be used to spot those whose scores do not do justice to their abilities because of an inappropriate attitude toward the tests or because of an unusual proneness to clerical errors. This is quite a large number of possible explanations for a single low score. Obviously the low score does not come with a label attached to tell us which is the right explanation in the given case. It is

expected, however, that reasonably accurate inferences can be drawn on the basis of the entire pattern of the student's scores on the rest of the tests.

7. Measures of scientific attitude. There are no clear historical antecedents for interspersing items designed to measure scientific attitude among items of an information test. This is one of the scales in the Project Talent test battery that are experimental in nature.

Description of the Talent Information Test

1. General characteristics. The Information Test of the Talent Battery consists of 395 items. All of these items, with the exception of the ten comprising the Scientific Attitude scale, are clearly "information items"; in other words, they test possession of factual knowledge. The ten scientific attitude items probably represent a composite of information, attitude, and judgment. Almost all of the 395 items are very brief.

As has been indicated, the Information Test was divided into two parts on the basis of whether subscores were to be punched on the scoring machine or specific item responses on the document reader. Sixteen scores were obtained initially for each student--15 subscores in Part I and their sum (the Part I total)--but potentially the test can yield a very large number of additional scores, since the items for which individual responses were punched (the Part II items) can be combined and recombined in any way desired and can also be combined in any way with any of the sixteen Part I scores. Thus scales can be tailor-made for specific purposes, such as predicting success in a particular occupational area. The tailor-made scale would combine whatever

items are best for the specific purpose.

Table VII-1 shows the number of items of each kind. The time limit allowed was as generous as administrative considerations would permit, so that almost everyone would have time to finish. Within the two parts the items were arranged in a roughly "spiral" order (i.e., the different kinds of items were intermingled apparently randomly but actually in a roughly cyclical arrangement). The arrangement of items was subject only to certain minor limitations in Part I, imposed by the requirements of the scoring machine and by the desirability of conforming at least approximately to *a priori* specifications concerning the specific locations of the Screening items. (The reasons for these *a priori* specifications are discussed later in this chapter.)

The following are a few sample items for this test that would fall in the miscellaneous category (category 38 in Table VII-1).

Item 9. In the song that starts "My country, 'tis of thee", the next line is

A. Of thee I sing.
B. Author of liberty.
C. Sweet land of liberty.
D. Protect us by Thy might.
E. Our fathers' God, to Thee.

(The answer is C.)

Item 10. Pulitzer Prizes are awarded for all of the following *except*

A. contributions to science.
B. biographies.
C. plays.
D. novels.
E. news photographs.

(The answer is A.)

Item 11. The best known British humor magazine is

A. What-Ho!
B. Last Laugh.
C. Cheers.
D. King's English.
E. Punch.

(The answer is E.)

On the basis of evidence obtained from the use of general information tests in the past, there is good reason to expect the total score on the Project Talent Information Test to prove useful as a predictor of college success.

Vocabulary Scale

This consists of items of the Information Test that provide a measure of the student's general vocabulary. This concept of "general vocabulary" probably requires explanation. Although the distinction between "general" vocabulary and "technical" or "specialized" vocabulary is not a precise one, general vocabulary can be considered to consist of words that might reasonably be in the vocabulary of well educated persons whether they have had specialized training in a particular area or not.

For instance, it would not be unreasonable to include items such as the following in the test:

Item 12. Which of these objects is usually spherical?

A. A water glass
B. An orange
C. A doughnut
D. A dime
E. A cone

(The answer is B.)

Item 13. The government of the United States is called "federal" because it has

A. a President.
B. three separate branches.
C. free elections.
D. separate states.
E. separation of Church and State. (The answer is D.)

Item 14. Another word meaning "vapor" is

- A. water.
- B. gas.
- C. liquid.
- D. condensation.
- E. solid.

(The answer is B.)

These three items test whether the student knows the meaning of "spherical", "federal", and "vapor". These terms represent concepts from the areas of mathematics, social studies, and physical science respectively, but they are nevertheless terms which it is reasonable to consider as part of a person's general vocabulary. Inclusion of items on words such as "spherical", "federal", and "vapor" is one way (a relatively minor way, perhaps, but nevertheless one that might have interesting results) in which the Talent Vocabulary scale differs from many widely used vocabulary tests, a number of which appear to be limited to words of a rather "literary" character. Though the Talent Vocabulary scale represents a wider range, note how words it might include, like "spherical", "federal", and "vapor", contrast with the following terms, which are far more highly specialized:

Mathematics terms: "binomial", "apothem"

Social studies terms: "pocket borough", "fief"

Physical science terms: "halogen", "ohm"

Items testing knowledge of the meanings of any of these six terms could legitimately go in the appropriate subject-matter scale (mathematics or social studies or physical science) but not in the Vocabulary scale. The reason that inclusion of specialized technical vocabulary would be undesirable is that in most cases knowledge of such terms

may depend far more on specialized training in a certain area than on general level of vocabulary.

Now let's get back to Sample Items 12, 13, and 14 above. It will be noted that not all of them are of the traditional vocabulary item form, calling for a synonym or definition. Item 14 does call for a synonym but Items 12 and 13 test comprehension of the words "spherical" and "federal" somewhat less directly. (If the student knows what "spherical" means it is reasonable to assume he also knows what an orange looks like and that he would therefore know that an orange is spherical. Similarly, if a student understands the concept represented by the word "federal" he can answer Item 13.) The recognition that comprehension of a word can be tested indirectly, that there is no particular merit in limiting the Vocabulary scale to items calling for synonyms or formal definitions, results in a better test. Its merit lies in the fact that it permits phrasing the items in whatever seems like the most promising way of determining whether the student understands the concept being tested and whether his degree of understanding is at a suitable level of specificity. (Deciding the appropriate level of specificity at which to pitch an item--the specific point to test for--is a matter which inevitably involves an element of judgment. In the case of the word "spherical", for instance, the point to test would be whether the student knows what something that is spherical looks like; since we are supposing it is to be regarded as a general vocabulary item, not a mathematics item, there would be no point at all in finding out whether the student knows the formal definition that a sphere is a surface formed by all points at a given

distance from a given point.)

The Vocabulary score is intended to measure ability to comprehend words. It is not intended to measure "recall" of vocabulary--i.e., the ability to select and use the appropriate word to express one's meaning. In line with the purpose of the test, to measure the ability to comprehend the writing of others rather than the ability to express oneself in writing, "understanding of a word meaning" is considered to mean understanding at a fairly gross level.

Very fine differentiations in meanings and implications, which are more a matter for the master stylist in English than for the "journeyman users" of the language, are not a matter of concern. Thus, although the dictionary indicates a slight difference in implication between "questionable" and "doubtful" (e.g., when used as the final word in the sentence: "His honesty is _____________."), the meanings are sufficiently similar that for purposes of this test the two words can be considered synonymous. For instance, the following item would be acceptable:

<u>Item 15</u>. "Questionable" means

A. inquiring.
B. doubtful.
C. inquisitive.
D. true.
E. false. (The answer is B.)

A further point, which perhaps may seem self-evident, is that the test is concerned with <u>written</u> English, not with spoken English. The fact that a person knows the answer to a certain question means that he understands the word when he encounters it in a written form, but this is no guarantee that he will understand it in its spoken form. There is a possibility that he won't, particularly if the word is one

which is pronounced quite differently from the way its spelling would suggest--"indict", for instance.

Scales to Measure Information in Specific Areas

In discussing those scales we shall divide them, purely as a matter of convenience, into "academic" subject matter areas and "non-academic" areas, but it should be borne in mind that this distinction, as it applies to the test, is a somewhat arbitrary one, since coverage includes both things learned in school and things that may have been learned outside of school, through independent reading, radio, television, and other available sources.

Academic areas. Under this heading are included those areas which constitute the major portion of the typical academic curriculum in high school.

The general purpose of these subscores is to measure acquisition of information in the area in question. An effort was made, therefore, to sample the field as broadly as was feasible with relatively few items.

1. Literature. The purpose of this scale is to measure familiarity with the world of literature. The chief "common body of literature" to which a sizable proportion of the group tested is likely to have been exposed--apart from the nursery rhymes and fairy tales of the preschool age--is the body of literature which is either required reading in school or is on school-issued lists of recommended reading.

No two lists of required or recommended reading issued by schools or school systems are identical. In fact it is

probable that any two such lists would differ considerably in detail--not only in regard to the grade level at which a certain book is placed but also in regard to whether certain books are included at all. This being the case, there is no one list or group of books, and probably not even any single book, that would be recommended or required reading for all students. However, a broad coverage makes it likely that students who have acquired the habit of reading will get reasonably good scores--at least if their reading is of materials above the level of comic books and is not limited to purely expository material such as "How to Build a Hi-Fi Set".

In developing the literature scale, a systematic effort was made to insure that both poetry and prose were well represented and further, that there was adequate representation of both children's literature (e.g., "Treasure Island", "Robin Hood", etc.), and literature that is considered suitable for adults as well as adolescents. Many of the items test knowledge of some salient point that would be found chiefly by <u>reading</u> the literary work in question rather than by merely reading <u>about</u> it. In this category are items about characteristics of principal characters, items about plot, etc. An effort was made, however, to avoid items about trivial or obscure points (minor incidents, unimportant characters, etc.). Items about specific works of literature are supplemented by items that test additional kinds of information (e.g., the meaning of such terms as "blank verse").

The following are sample items:

Item 16. Each line of blank verse has

A. 6 feet.
B. 10 feet.
C. 3 feet.
D. 4 feet.
E. 5 feet.

(The answer is E.)

Item 17. Othello's motive for murder was

A. self-defense.
B. jealousy.
C. fear.
D. craving for power.
E. greed for gold.

(The answer is B.)

Item 18. Which of these was a knight of King Arthur's Round Table?

A. Alfred
B. Lancelot
C. Merlin
D. Ivanhoe
E. Roland

(The answer is B.)

2) Social studies. This scale covers facts and concepts from the fields of history, economics, government and civics, geography, current affairs, etc. At least half of the history, government, and geography items are concerned with the United States, but there is coverage of the rest of the world too. Social studies content taught in elementary school has been sampled, as well as that taught in high school.

The sampling of material on United States history and government includes some coverage of periods from the discovery of America to the present. Political, economic, industrial, and military history were sampled. In regard to the coverage of non-U.S. history, the items are chiefly concerned with the history of the Western world. It was felt

that the history of the Orient--at least its history in regard to matters which have not impinged directly on the West, and certainly its history in the eras preceding its opening to Western influence--is so little taught and so little known that items on such matters would be too difficult to contribute a worthwhile degree of reliable information to the test scores.

Items that are likely to become dated within 20 years or whose significance is likely to change markedly within that period were kept to a minimum. However, this did not wholly preclude items on topics of current concern. Some situations which have recently been very much in the public eye will probably be matters of continuing interest and concern for a long time.

Both vocabulary items and items testing other kinds of information are included.

Materials not necessarily taught in high school are not rigidly excluded from the test, since it is felt that acquisition of knowledge about such materials through voluntary reading or in other ways indicates special interest in the area. The following are sample items:

<u>Item 19</u>. Which of these is specifically required by the U. S. Constitution?

A. Supreme Court
B. Political parties
C. Congressional committees
D. Nominating conventions
E. Cabinet

(The answer is A.)

Item 20. Which of these countries has more than one native language?

A. France
B. Greece
C. Italy
D. Switzerland
E. United States

(The answer is D.)

Item 21. Napoleon was emperor of France around the beginning of the

A. 15th century.
B. 16th century.
C. 17th century.
D. 18th century.
E. 19th century.

(The answer is E.)

3) Mathematics. The area of mathematics, unlike other academic subject matter areas, is represented in the Talent battery not only in the Information Test but in the following other tests as well:

a) Arithmetic Computation Test

b) Mathematics Test

Part I.	Arithmetic Reasoning
Part II.	Introductory
Part III.	Advanced

However, the mathematics items in the Information Test do not overlap the content of these other tests since the Information Test items are concerned with knowledge of facts, not with the possession of mathematics skills. No Information Test items require computation. (This is covered in the Arithmetic Computation Test.) No Information Test items require reasoning or problem-solving. (This is covered in the Mathematics Test.) Nearly half of the items are on topics covered in Grade 9 or earlier, and the others are concerned with topics that are not usually touched upon until at least the 10th grade.

Some items on information that might be acquired in other ways than through formal instruction in mathematics were included since it was felt that such information might indicate special interest in mathematics. (An example of such information, not generally included in the high school mathematics curriculum, is the impossibility of trisecting an angle by Euclidean means.) The history of mathematics was not included, however.

Within the limitations imposed above, the items sampled the field as broadly as seemed feasible.

Both vocabulary items and items testing other kinds of information were included. Among the kinds of content sampled were such matters as definitions, the vocabulary of mathematics, conventions of mathematical notation, and the understanding of concepts (e.g., fractions, decimals, percentage, factors, equations, planes, angles, logarithms, exponents, etc.).

Sample items follow:

Item 22. "$3b + 2b^2 = 5$" is called

A. a coefficient.
B. an equation.
C. a formula.
D. an unknown.
E. an identity.

(The answer is B.)

Item 23. An ordinary slide rule is designed for

A. computing.
B. measuring length.
C. measuring angles.
D. weighing.
E. reading tables.

(The answer is A.)

4. Physical science. This scale includes items from the areas of chemistry, physics, astronomy, and other physical sciences.

Science topics taught in elementary school are included, as well as those not studied until high school. Both vocabulary items and items testing other kinds of information were included. Items on information that might readily be acquired in other ways than through formal instruction in science were included since it was felt that acquisition of such knowledge through voluntary reading, experimentation, or in other ways would indicate special interest in the area. Sample items follow:

Item 24. Ozone is a form of

A. electricity.
B. air.
C. oxygen.
D. nitrogen.
E. carbon dioxide.

(The answer is C.)

Item 25. Gravitation is the force which causes a rubber ball to

A. bounce.
B. go flat.
C. keep its shape.
D. fall when dropped.
E. change shape when pressed.

(The answer is D.)

Item 26. A "light year" is a unit of

A. temperature.
B. volume.
C. brightness.
D. distance.
E. speed.

(The answer is D.)

5. Biological science. This scale includes items from the fields of botany, zoology, and microbiology.

Items about nature lore are included, as well as items concerned with the more formal aspects of biological science.

Both vocabulary items and items testing other kinds of information are included.

A sample item follows:

Item 27. The substance which makes plants green is

A. cellulose.
B. starch.
C. chlorophyll.
D. sap.
E. fungus.

(The answer is C.)

Other Areas

Reference to Table VII-1 shows the many kinds of information covered, besides those in areas taught as part of the typical academic curriculum in high schools.

In developing these items, an effort was made to include some that would give an advantage to those who had participated in the activity in question as a hobby or voluntary extra-curricular activity (e.g., electronics, carpentry, mechanics, cooking, sewing, photography, sports, theater, ham radio, music, art, journalism, chess, other games, etc.). Other items were designed to give an advantage to those who had participated in the activity as a household chore (e.g., cooking, sewing, baby care, home

maintenance and minor repairs, farm chores). Still other items were intended to give an advantage to those who had participated in the activity in question through a part-time or summer job (e.g., baby-sitting, office jobs, etc.). Some items give an advantage to those who have done extracurricular reading which has a direct or indirect bearing on the area in question (e.g., engineering, architecture, law, medicine, aeronautics and space, photography). Items are also included that give an advantage to those who have had some other form of special training or education.

Some of the characteristics of specific scales or subsets of items follow.*

1. Music. While these items certainly are not indicative of musical talent, those who enjoy going to concerts and opera or listening to music on the radio and phonograph should have an advantage on this set of items (except perhaps if their interest is strictly limited to rock-and-roll).

Here is a sample item:

Item 28. Who wrote "Peter and the Wolf"?

A. Prokofieff
B. Bach
C. Tschaikowsky
D. Dukas
E. Mozart

(The answer is A.)

2. Art. This scale measures knowledge about art. This knowledge may be acquired through participating in art as a

---

*Note that three scales, Aeronautics and Space, Electricity and Electronics, and Mechanics, because of their major importance, are discussed separately, following the discussion of the Information Test.

hobby or recreational activity or in other ways. Those who enjoy going to museums and art galleries, reading about art, and perhaps dabbling a little in it themselves should have an advantage on these items--as should those who have had special training in art. However, there is absolutely no reason to think that this scale will identify those with creative talents in art.

Two sample items follow:

Item 29. Easels are used by

A. sculptors.
B. musicians.
C. surgeons.
D. lawyers.
E. painters.

(The answer is E.)

Item 30. Which of these is not the name of a famous painter?

A. Titian
B. Rubens
C. Cezanne
D. Napoli
E. Van Dyck

(The answer is D.)

3. Home Economics. These items should give an advantage to students who have engaged in relevant activities (cooking, sewing, etc.) as a household chore or because they enjoy them, those who have engaged in relevant activities as a part-time job (e.g., baby-sitting), and those who have taken home economics courses.

A sample item follows:

Item 31. In a recipe calling for one cup of butter, you should use

A. one pound.
B. two pounds.
C. a half-pound.
D. a quarter-pound.
E. one-eighth pound.

(The answer is C.)

4. Law. While these items primarily measure a kind of "general information", so that a good score on this scale is not necessarily indicative of particular interest in this field, it nevertheless seems likely that a student whose score in this area is relatively low is not likely to have enough interest in law to make it an advisable career for him.

Here are two sample items:

Item 32. What kind of degree do most lawyers have from law school?

A. B.A.
B. B.S.
C. Ph.D.
D. LL.D.
E. LL.B.

(The answer is E.)

Item 33. A coroner's chief function is to investigate

A. mysterious deaths.
B. burglaries.
C. counterfeiting rings.
D. kidnapping cases.
E. accusations of bribery.

(The answer is A.)

5. Medicine. Areas sampled include knowledge of physiology, nutrition, first aid, common ailments, physicians' instruments (e.g., stethoscope), training requirements for careers, etc.

Here is a sample item:

Item 34. A deficiency disease is a disease caused by

A. a virus.
B. improper diet.
C. not getting inoculated.
D. defective glands.
E. poor heredity.

(The answer is B.)

6. Architecture. Those who have read books and articles about architecture, and students whose interest has manifested

itself in other ways--for instance, those who have been unusually observant in regard to the architecture all around them--should do well on these items.

7. Farming. These items are intended to give some representation to the kinds of special information that children who grow up on farms or ranches are likely to acquire. The items might give some advantage to those who have an interest in these aspects of rural life.

8. Theater. These items should be indicative of an interest in the theater. (It is somewhat doubtful, however, that it is possible to get much differentiation between those who are interested in the stage as a career and those who merely are interested as spectators.) Most of the items deal with theater terms (e.g., grease-paint, footlights, etc.).

Here is a sample item:

Item 35. An ingenue role in a play usually calls for:

A. exceptional acting ability.
B. great beauty.
C. youthfulness.
D. a sophisticated manner.
E. skill as a comedian.

(The answer is C.)

9. Photography. These items should test information that is acquired principally through having photography as a hobby, taking pictures, processing them, and also, to a certain extent, through reading about photography, attending exhibits, and manifesting interest in other ways.

Screening Scale

The primary purpose of the Screening items is to identify those students whose reading disability is so

severe that their scores on certain other tests are automatically invalidated.

Certain tests of the battery in addition to the Reading Comprehension test have a high enough verbal content that scores on them will be affected by gross reading disability. (Among these tests are Information, Arithmetic Reasoning, Mechanical Reasoning, Creativity, and Table Reading.) The three inventories (Student Activities Inventory, Interest Inventory, and Student Information Blank) will also be affected. The primary purpose of the Screening Scale in the Information Test is to identify students whose scores on the tests named are questionable because of lack of ability to read well enough to comprehend the questions. A secondary purpose is to identify those students who, though they can read adequately well, do not obtain scores which reflect their true ability, either because of a tendency to make numerous careless mistakes of a clerical nature (such as accidentally marking an answer space other than the intended one) or because of an unfavorable attitude toward the test-taking situation. An unfavorable attitude--reflecting faulty motivation in regard to the battery--might be manifested in the form of either apathy or flippancy. (It should be noted, however, that some students might make the flippant, deliberately incorrect response only in the case of the screening items and other extremely easy items.)

Those with a low screening score will fall into several categories, as follows:

a. Non-readers

(1) By far the largest category will probably be the low IQ group--those who can't read because of lack of

normal intelligence. (These examinees will probably get low scores on all the tests, not just the ones mentioned above as being directly dependent on reading ability.)

(2) A second and much smaller category will consist of those whose chief difficulty has been in learning to read and who appear to have normal intelligence in other respects. Such examinees, theoretically, might score quite well on some of the less verbal tests in the battery--e.g., Object Inspection, Visualization in Two Dimensions, Visualization in Three Dimensions, Abstract Reasoning.

(3) There may also be some foreign students who haven't learned enough English yet to be able to cope with the verbal tests. (These students should be identifiable from the Student Information Blank and other auxiliary sources of information.)

(4) Some instances of reading disability may stem from the fact that the student has been deprived of a normal education--by illness or physical disability, for instance. (These students, too, may be identifiable from the Student Information Blank and other auxiliary sources of information.)

b. Others, whose low scores are primarily reflections of poor motivation in the test-taking situation

(1) A few may make numerous clerical errors (accidentally marking answer spaces other than those corresponding to the intended answer). These students should be characterized by a very low score on the Clerical Checking Test and the Table Reading Test, which are measures of clerical speed and accuracy.

(2) Those whose general attitude is apathetic

or indifferent can be expected to have many low scores in the less verbal tests as well as in those requiring good reading ability.

(3) Those whose attitude is flippant can be expected to have a peculiar pattern of item responses--for instance a pattern which yields a score much lower than a pattern of purely chance responses would be expected to give.

(4) There may be some whose attitude is flippant only in regard to certain kinds of items. For instance, the extreme easiness of the Screening items may in itself induce in some students a flippant attitude towards these items--manifested by the deliberate choice of wrong answers. Such behavior can be spotted by the fact that the Screening score would be much lower than would be reasonable in view of the pattern of the other Information scores.

There is no expectation whatever that it will be possible on the basis of the Screening score alone to differentiate among the several categories of "non-readers" and badly motivated students mentioned above. But the Screening score will differentiate those who fall in any of these categories from students who have normal reading ability and are motivated to do well. And the pattern of other test scores, as suggested above, may help to differentiate among the various categories. A low Screening score will indicate that in the absence of evidence to the contrary (e.g., low scores on all tests--even those not requiring reading), scores on Arithmetic Reasoning, Mechanical Reasoning, Creativity, the rest of the Information scales, etc., and the responses on the various inventories may be quite misleading. In other words, it is conceivable, though not necessarily actually true, that

the examinee's scores on these tests might be much higher if someone were reading the test items to him, and recording his responses for him.

To serve these purposes the items are extremely simple, involving concepts that almost everybody above the level of a low-grade moron would be familiar with.

The vocabulary has been kept as simple as is reasonable (as is also the case for the tests whose scores may be inferred to be invalid for those with low Screening scores. In the case of these latter tests, it may be remarked parenthetically that the vocabulary should be very simple except where more difficult words are necessary in order to test whether the student comprehends a certain concept.)

The Screening items have been arranged in the Information Test in such a way that they occur very frequently at the beginning of the test, then less and less frequently so that the intervals between successive Screening items increase in accordance with a rough geometric progression, and then finally a few items are concentrated at the end. The purpose of this arrangement is to keep the students that are having difficulty with the test aware that very easy items occur sporadically. The aim is to encourage these students to read every item.

The following is a sample Screening item:

Item 36. How many days are there in a week?

A. 2
B. 3
C. 5
D. 7
E. 10

(The answer is 7.)

Scientific Attitude Scale

These items provide a subscore which should be indicative of how the individual views the world--whether, on the one hand, he views it as a place where there are logical cause-and-effect relationships or whether, on the other hand, he regards it as a place where consequences are illogical, capricious, and arbitrary. In the latter category (illogical and arbitrary consequences) fall the modern-day equivalents of the kinds of primitive beliefs that Frazer*, in his epochal treatise on cultural anthropology, The Golden Bough, terms "sympathetic magic". Also in the category of illogical and arbitrary consequences are the premises of palmistry, astrology, numerology, and occult manifestations of all sorts. Among the students who should get low scores on this scale are:

a. Those who are superstitious, in the sense that they believe such things as that black cats bring bad luck, horseshoes bring good luck, Friday the 13th means trouble, there is something in astrology (at least when the astrologer is an expert), etc.

b. Those who believe that certain people--e.g., gypsies, mediums, etc.--have occult powers of one sort or another--for instance that they can see into the future or converse with the dead.

c. Those who favor anthropomorphic explanations of the behavior of lower animals.

d. Those who believe that the world is full of evil

---

*Frazer, J. G., The Golden Bough, New York, Macmillan, 1922.

spirits that have to be propitiated.

e. Those who do not seek a logical explanation for unexplained phenomena, but instead tend to accept naively the first explanation presented that is superficially attractive.

Among the people who should get high scores are those who, like the scientist, seek reasonable and logical explanations for events and unexplained phenomena, and tend to reject explanations that involve occult arts, magic, etc. It should be noted that a person may have a scientific attitude even if he has never studied science formally and has little or no factual information about how scientific research is actually carried out. It will be seen that the trait to be differentiated by these scores involves not only intellectual ability (the judgment and reasoning power necessary to decide whether a "conclusion" follows logically from given premises) but also a basic tendency to use the scientific approach.

Each item consists of a description of a phenomenon or an occurrence. Five explanations are presented, and the task is to select the "best" one. Only one of the five explanations is reasonable and logical. Of the other four, some involve common superstitions, magic, concepts incompatible with the scientific viewpoint, belief in the occult, etc. Selection of other distractors, however, may primarily imply muddled thinking on the part of the examinee. (An example of the latter kind of distractor would be Option C in Sample Item 37, below.) One reason for including such options is that it is probably undesirable to have all distractors fall in the other category (superstition, the

occult, etc.) since this makes the items too obvious.

None of the items requires possession of any special information (in science, mathematics, etc.). Whenever such information is needed to answer the question, it is presented in the stem of the item. For instance, in Sample Item 38 below, the probability that ten coins will fall "heads up" is stated, in round numbers, in the item.

<u>Item 37</u>. Jim Wilson has entered many golf tournaments but has never won one. He usually does very well until the finals. But in the finals when the score is close, he tends to make a few wild shots, which result in his defeat. The best explanation of Wilson's failure to win a tournament is that

A. he is just naturally unlucky.
B. he "goes to pieces" under pressure.
C. he hasn't practiced.
D. he doesn't really care whether he wins.
E. his opponents are just naturally lucky.

(The answer is B.)

<u>Item 38</u>. Professor Rogers wished to find out whether any of the 950 students in Central High School could demonstrate the power of "mind over matter". When ten pennies are tossed, the chances that all ten of them will fall "heads up" are about one in a thousand. Rogers had each student in turn toss ten pennies. He instructed them to try, by thinking very hard about it, to make all ten pennies fall "heads up". But when one of the boys, Joe Thompson, tossed the coins they all fell "tails up". What does this suggest about Joe?

A. Joe was purposely trying to get all tails.
B. Joe became confused.
C. Joe didn't have faith in the power of mind over matter.
D. Joe is unlucky.
E. Nothing.

(The answer is E.)

It is hypothesized that scores on this test are related to science aptitude.

Mechanical Information Scale*

Purpose of scale. Mechanical Information Scales are usually intended to provide a measure of aptitude for jobs in the mechanical area. The underlying assumption is that anyone who has aptitude in this area and also has sufficient interest in it that he would enjoy this type of activity would manage in one way or another to acquire considerable information in the area. It is assumed that he would be likely to acquire this information whether or not the high school he attended happened to offer formal instruction in mechanical subjects.

Historical antecedents. Among the first standardized tests of mechanical information were the early trade tests developed for use in the Army during World War I. These were specific information tests developed for various sorts of mechanical occupations, and their purpose was to measure the achievement of individuals in those fields. The items were of types much like those in current use. The Army Alpha too had a few information items on mechanical topics. In the early tests of mechanical information, unlike the more recent ones, no distinction was made between mechanical information (tools, materials, mechanics, carpentry, cabinet making, painting, printing, surveying, etc.) and electrical information. In current practice electrical and electronic information are often covered in a separate test since individuals who do well in mechanical information tests are not necessarily

---

*As has already been indicated, these items are embedded in the Information Test.

well informed in the areas of electricity and electronics, and vice versa. Contemporary tests of mechanical information have proved useful for predicting success in many types of jobs, including aircraft mechanic, auto mechanic, radio repairman, radar repairman, machinist, sheet metal worker, welder, etc. In World War II, the Army, Navy, and Air Force all made extensive use of mechanical information tests with considerable success.

In part, mechanical information tests measure reading ability and general learning ability (the ability to assimilate information when exposed to it); in other words, abstract intelligence. However, in a well designed mechanical information test the demands on reading skills are minimal even if verbal rather than pictorial items are used. An effort is made to keep the wording of the items simple.

Pronounced sex differences are usually found in scores on mechanical information tests; these differences, however, are often less pronounced when the items are of the verbal type than when they are pictorial.

In summary, mechanical information tests chiefly provide measures of what has sometimes been called the "mechanical experience factor".

Description of Talent Mechanical Information Scale. The items in this scale are all of the verbal type, not the pictorial type. Each of the items falls into one of the following two categories:

1. Information about tools, carpentry, and other kinds of construction

2. Information about motors and other kinds of mechanisms

The items are about equally divided between the two categories. As it happens, however, this is not a particularly important distinction (except as a way of insuring breadth of coverage), since those who do well in one of the two categories tend to do well in both; in other words, scores on the two sets of items are quite highly correlated. For this reason the two categories of items have been combined in a single scale.

The test taps a wide range of mechanical information. Many of the items call for information about automobiles; other items are concerned with other common machines and tools with which boys who are interested in mechanical activities are likely to be familiar. The emphasis is on information acquired through <u>direct</u> experience with tools and motors. The very wide variety of information covered tends to equalize the opportunities of boys in urban and rural environments to score well and helps make the scores comparable.

A sample item follows:

> <u>Item 39</u>. Which of these tools is best for rough-shaping wood?
>
> A. Rasp
> B. Plane
> C. Level
> D. Awl
> E. File
>
> (The answer is A.)

Having briefly described the kinds of items the Mechanical Information scale includes, perhaps it would be well to say a word or two about what it does <u>not</u> include. Items to measure information that would be acquired in an academic course in science (e.g., a course in physics) are

included in the Physical Science Information scale rather than in the Mechanical Information scale except in the case of items which primarily stress a practical application relevant to mechanics. Neither scale has items requiring comprehension of mechanical principles and everyday physical forces (e.g., gravity, equilibrium, pressure, etc.). Such items are in another Talent test--the Mechanical Reasoning Test. Nor are items requiring a comprehension of electricity included in the Mechanical Information scale even if they are concerned with motors. Such items are in the Electricity and Electronics scale.

In interpreting scores on this test, past mechanical experience or formal training in mechanics should of course be considered. Certainly, too, the sex of the individual should be borne in mind in interpreting a score. In general those who score well in Mechanical Information will tend to do well in training for a wide variety of mechanical occupations; for instance, machinist, artillery mechanic, sheet metal worker, foundryman, aircraft mechanic, auto mechanic, Gunner's Mate, and aircraft pilot.

Electrical and Electronic Information Scale*

Purpose of scale. This scale is intended to serve as a predictor of success in jobs in the area of electricity and electronics (e.g., electrician, electronic technician, radio and television repairman).

---

*As has already been indicated, these items are embedded in the Information Test.

Historical antecedents. There is nothing particularly novel about using electrical and electronic information tests with content geared to a specific training course as straightforward measures of achievement in the training program. But extensive use of electrical and electronic information tests with more general content as predictors of success in relevant kinds of jobs does not seem to date back any further than World War II.

At that time, the military services found that tests of electrical and electronic information were useful in identifying persons who would make good electricians or electronic technicians (though not necessarily good mechanics). This sort of test was therefore used extensively and successfully in assigning military personnel to specific training programs. Ever since then it has been realized that the occupational fields represented by mechanical jobs and electrical jobs are functionally different. Civilian versions of electrical and electronic information tests for use in selection and counseling were not developed until quite recently ( ).

Electronic information tests have proved useful to predict success in numerous jobs, including electrician, electronic technician, radio repairman, radar repairman.

A word of caution may be advisable at this point. When tests are being used for selection it is perhaps even truer in the case of jobs in the electrical and electronics field than in other fields that decisions should not be based on just one test. The entire pattern of scores on several other aptitude and achievement tests is also highly relevant in the case of most such jobs.

Description of Talent Electricity and Electronics Scale. Most of the items in this scale are of the verbal type, though a few do involve diagrams. The items stress information that is acquirable through direct experience in the construction and maintenance of electrical and electronic equipment. Thus, for instance, examinees who have worked on radios, hi-fi sets, or other electronic equipment would have an advantage. It is inevitable that some of the items will also be answerable on the basis of information acquired in formal courses in physics and to a lesser extent in chemistry. However, items that can be answered on this basis even by persons who have not had much direct experience in constructing or maintaining electrical or electronic equipment should be kept to a minimum.

Scores on the Electricity and Electronics scale undeniably are considerably affected by past experience in the area, but since a broad range of content is covered, the average high school student will probably have had an opportunity to acquire a fair amount of the information tested if he is at all interested in this area and has any aptitude for it. In other words, those who get good scores on this scale are likely to have both interest in the field and aptitude for it (as indicated by the demonstrated ability to assimilate this type of information). They could probably successfully complete any training course in the area. A low score on the scale does not indicate that a youngster lacks the ability to learn this kind of material if he is exposed to it. However, if he has not been exposed to this area by age 16 or 18 and has no knowledge of it, the prognosis for success in an electronics or electrical

training program would be substantially lowered.

Aeronautics and Space Scale*

Purpose of scale. This scale is intended to help identify students whose interests and aptitudes suggest that they would be successful pilots.

Historical antecedents. As has already been indicated, among the discoveries made by the Aviation Psychology Research Program during World War II was the fact that one of the tests that made an important contribution to the prediction of success in a pilot training program was an information test heavily loaded with questions about aeronautics and related topics.

---

*As has already been indicated, these items are embedded in the Information Test.

Description of Talent Aeronautics and Space Scale. The items are on such topics as piloting procedures, navigation, jet planes, space exploration, etc.

B. Memory for Sentences Test

Purpose of test. The purpose of the test is to measure one kind of memory--the ability to memorize and recall meaningful material.

Historical antecedents. This test is experimental in nature. Numerous tests of ability to memorize materials have been developed in the past, but as far as the authors of this book know, the type of item used in the Talent Memory for Sentences Test has not been used before in precisely the same form. This item type is a multiple-choice

version of a kind of completion test that had been developed to measure essentially the same thing.

This earlier test consists of a study period in which a number of sentences are to be memorized, followed by a testing period in which some or all of the same sentences are presented in a different order, with one word missing from each. The task is to fill in the missing word.

Description of Talent Memory for Sentences Test. This test consists of about 40 short sentences, which the students are given six minutes to memorize. The test items consist of some of these sentences, in a different order, with one word omitted from each sentence. The task, which the examinees are told about at the time they are asked to study the sentences, is to recall the missing word and identify its second letter, which is one of five options. Each option consists of only one letter, instead of an entire word, so as to test the ability to actually recall the missing word, rather than merely the ability to recognize the word when it is presented. The second letter is used instead of the first in order to minimize the extent to which the letter itself will help the person to recall the missing word if he has forgotten it.

There are several different dimensions in terms of which memory tasks can be classified:

1. Recall vs. recognition
2. Short-term memory vs. long-term memory
3. Incidental memory vs. deliberate memorizing
4. Memory for "underlearned" materials vs. memory for "overlearned" materials
5. Kind of material to be remembered (e.g., verbal material, diagrams or pictures, music)

6. Meaningfulness of materials to be remembered (e.g., meaningful material such as paragraphs vs. meaningless material such as nonsense syllables).

7. When the material to be learned is meaningful, what is the nature of the learning task? Must the material (e.g., sentences) be memorized verbatim or is it sufficient merely to remember the sense and be able to paraphrase it?

Where does the Talent Memory for Sentences Test stand in respect to these seven dimensions?

In the first place we believe it measures recall rather than recognition. That point has already been discussed but perhaps it should be underlined for the benefit of those critics of objective tests who say that it is impossible to measure recall with an objective test--that all that can be done is to measure recognition. The authors disagree, and offer the Memory for Sentences Test as evidence.

In regard to the short-term long-term dimension, the test measures relatively short-term memory. However, because two other tests (Arithmetic Computation and Memory for Words) intervene between the study period and the test period of the Memory for Sentences Test, the interval between study and test is sufficiently long (about 20 minutes) that something more than immediate memory is tested. Furthermore, the task is made considerably more difficult by the retroactive inhibition (if we may lapse into psychologist jargon for a moment) set up by the intervening memory task (memory for words).

Memory for materials which the student deliberately tries to memorize is measured--not incidental memory.

Memory for materials which are probably underlearned is measured--not memory for overlearned materials.

The material to be memorized is verbal (i.e., sentences).

The material is moderately but perhaps not maximally meaningful. It is moderately meaningful because it consists of meaningful sentences- but they would be much more meaningful if they were not isolated sentences, completely removed from context.

The task is such that verbatim memorization is required. Synonyms for the missing words are not acceptable. (In this connection it should be noted that items were not used where an obivous synonym of the missing word would have the same second letter.)

In the preceding paragraphs we have gone into considerable detail concerning what kind of memory is measured by this test because it is so easy (and so risky) for test users to overgeneralize from the title of a test and to infer, for instance, that because a test has the word "memory" in its title it will indicate how well someone can do in all kinds of memory tasks. Nothing could be further from the truth. There are many specific kinds of memory, and if results derived from a test of one kind of memory are generalized to other kinds it should be done only with extreme caution. Interpretation of results should be based insofar as possible on empirical data--for instance, specific studies of what kinds of variables the score on the test correlate with.

Having stressed that warning, perhaps it is safe at this point to mention that it is usually considered reasonable to regard the kind of memory measured by this test as one aspect of the complex of abilities subsumed under the term "general intelligence". Those who do well on the test are probably likely to do well in types of school work that call for rote learning of verbal materials,

and they will also probably make good initial progress in training programs where a lot of momenclature and specific facts are to be learned.

To give an idea of the character of the sentence to be memorized and the items to be answered, a sample of each is given.

The following sentence might have been (but wasn't) used as one of the 40 to be memorized:

"Mary tried to catch the four o'clock bus."

If that sentence had been used, the following could have been one of the test items:

Item 40. Mary _____ to catch the four o'clock bus.

A. _a _____ R. _r _____
E. _e _____ U. _u _____
I. _i _____ (The answer is R--the second letter of "tried".)

A systematic effort was made to insure that if a person knew the answer to an item he knew it because he had succeeded in memorizing the sentence, and not for any other reason. For this reason the four letters that were used as distractors were always the second letters of words that might reasonably fill the blank space in the sentence. For instance in the case of the sample item above, the missing word could conceivably be:

ran, wanted, had, failed (second letter: a)
needed, decided, went (second letter: e)
liked, wished (second letter: i)
hurried, rushed, (second letter: u)

To avoid confusion, the first letter of the missing word was never used as a distractor. Note, for instance, that in the sample item above, T (the first letter of "tried") was not used even though it was also the second letter of a plausible, though incorrect word ("attempted").

Likewise an effort was made to insure that if someone answered an item incorrectly it was because he did not

remember the sentence--not because of faulty spelling. Therefore, no item was used where the missing word was a hard one to spell, and for which, therefore, there might be some confusion or misconception in the mind of the examinee as to what the second letter is. For instance the following item would be a bad one:

"Dr. Phwimphsniggitch is a professor of psychology."

If "psychology" were the missing word in this sentence, and if the options were E, H, I, S, and Y, an examinee might choose Option H under the very prevalent misconception that "psychology" is spelled "phychology"; other ways of misspelling could easily yield I and Y as answers.

C. Memory for Words Test

Purpose of test. The purpose of this test is to measure one kind of rote memory--the ability to memorize the "foreign" words corresponding to common English words. This is an ability which is obviously directly relevant to the learning of a foreign language. It is also presumably related to the ability to learn many other kinds of material.

Historical antecedents. Various kinds of "paired associates" tests have been used for a long time in psychological research on the nature of memory, and also as components of intelligence tests and predictors of success in occupations where this kind of memory seems relevant. A "paired associates" test is a test in which a set of pairs of symbols is studied for a fixed period, to associate each symbol with the one paired with it. The symbols used may be of any nature (real words, nonsense syllables, numbers, diagrams, pictures, persons' names, etc.)

As early as World War I, paired associates were used to measure code learning ability. More recently, Carroll and Sapon (3) developed a "Paired Associates" Test (Test 5 in the Psi-Lambda Foreign Language Aptitude Battery) for use in predicting ability to learn foreign languages. This test has been shown to be useful for its intended purpose. Using a factor analysis technique* Carroll (2) found, for instance, that his "Paired Associates" Test measured a factor he designated as "associative memory" and to a somewhat lesser extent a factor he called "linguistic interest", both of which were related to course grades in an intensive Air Force "trial course" in Mandarin Chinese.

<u>Description of Talent Memory for Words Test</u>. This test is modeled directly on Test 5 (Paired Associates) of the Psi-Lambda Foreign Language Aptitude Battery.

In this test the student is given two minutes to study 24 common English words, and their alleged equivalents in a strange and totally unfamiliar language. To insure that the language would be unfamiliar to all, a language named Vlaznoor, invented specially for use in the Talent test, was used. The two-minute study period is followed by a two-minute practice period in which the students practice recall, using a list of the Vlaznoor words arranged in a different order. Ability to select the English equivalent of a given foreign (Vlaznoor) word from among five

---

*Factor analysis is a statistical procedure for analyzing the interrelationships among a set of variables.

options is then tested. Each Vlaznoor word is a separate item, and they are in a different order from those used in the learning and practice materials. Almost all options are English words used in the vocabulary studied, although an occasional English word not in the vocabulary studied may be used as a distractor--e.g., "in", when "on" is the right answer.

The foreign words are grossly different from each other. Words that differ in only one letter (e.g., DRINVO and DRANVO) are not used. Each of the foreign words is a pronounceable combination of vowels and consonants (e.g., KAJELB), not an unpronounceable jumble consisting solely of consonants (e.g., BTFLSPKT). The foreign words are mostly monosyllables, or at most disyllables.

A sample of the type of vocabulary used is shown below:

Foreign	English
KAJELB	walk
DRINVO	potato
HOL	ear
'	'
'	'
TAHNE	pretty
'	'
'	'
FALG	shoe

The following is a sample test item:

Item 41. DRINVO

A. ear
B. pretty
C. shoe
D. potato
E. walk

(The answer is D.)

This test measures ability to memorize pairs of arbitrarily associated expressions, when one member of the

pair is a meaningful English word, and the other is a meaningless verbal symbol.

This test differs from the other memorization test in the battery, the Memory for Sentences Test, in several major respects besides the very obvious one of the nature of the materials to be memorized.

The Memory for Words Test, unlike the Memory for Sentences Test, measures the recognition type of memory, not the recall type.

It measures immediate memory. There are no other tests intervening between the study period and the test period, as there are in the case of the Memory for Sentences Test.

The material to be learned is somewhat less meaningful than that of the Memory for Sentences Test, since half of each pair of terms to be associated is an arbitrary verbal symbol.

As for the similarities between the Memory for Words Test and the Memory for Sentences Test, both test deliberate memorization, not incidental memory; and both measure memory for somewhat "underlearned" materials rather than for "overlearned" materials.

There is considerable presumptive evidence, inherent in the nature of the task, that the ability measured by the Memory for Words Test is one that is very useful in learning a foreign language. In other words, the test has excellent "face validity" as a measure of aptitude for foreign languages. As has been indicated, this face validity is bolstered by considerable empirical evidence (obtained chiefly in connection with the Paired Associates Test of the Psi Lambda battery).

There is also good presumptive evidence that the test would be a useful predictor of ability to learn many kinds of code. (After all, the task imposed by the test is a form of code learning.)

Rote memory is also obviously useful in many other kinds of occupational endeavor.

D. Disguised Words Test

Purpose of test. This test is designed to measure what has been termed "phonetic-orthographic ability--the ability to form connections between letters and sounds" (2).

Historical antecedents. A test in which the task was to decipher words spelled in a partially phonetic and very much abbreviated form was used about 20 years ago as a part of the Turse Shorthand Aptitude Test (12). This test was demonstrated to be useful for its intended purpose of predicting ability to learn shorthand.

The task described above used in the Turse Shorthand Aptitude Test, was converted to a slightly more complex task which could be measured by multiple-choice items, in the Spelling Clues Test (Test 3 of the Psi-Lambda Foreign Language Aptitude Battery). In this version the task was to figure out what word was represented by the odd combination of letters and then to demonstrate that one had translated correctly by picking the correct synonym from among five choices. Carroll obtained empirical evidence that his Spelling Clues Test was useful as a predictor of foreign language learning ability.

Description of Talent Disguised Words Test. This test uses the same type of item as the Spelling Clues Test

described above, with the minor difference that in the case of the Talent Test a systematic effort was made to reduce the vocabulary level of the words "spelled" by limiting these words to approximately the first 5,000 in the "Teacher's Word Book of 30,000 Words",* which shows the order of frequency with which words are used in written English. The reason for imposing this restriction was to reduce to a minimum the extent to which the test measures vocabulary level, since there was already a separate vocabulary scale in the battery (incorporated in the Information Test).

---

*Thorndike, R. L., and Lorge, I. <u>The teacher's word book of 30,000 words</u>. New York. Teachers College, Columbia University, 1944. (The words used were limited to those designated AA, A, or 14-49 in the G column of the word list.)

The directions for the test read as follows:

> "Each item begins with a word which is disguised by spelling it in a very peculiar way. However, it is spelled somewhat as it is pronounced. The disguised word is followed by five real words or phrases.....Show that you recognize the disguised word by selecting the choice that means most nearly the same thing......"

Two sample items follow:

<u>Item 42</u>. DSKRIJD

A. depressed
B. described
C. thrown away
D. ruined
E. accused

(The answer is A, since the disguised word is "discouraged".)

Item 43. DLA

A. sadly
B. postpone
C. bluntly
D. hand out
E. every day

(The answer is B, since the disguised word is "delay".)

Perhaps citing some of the rules that had to be conformed to in writing the items will give a clearer idea of the character of the test.

The following were the only acceptable ways of distorting the spelling:

1. Replacing non-phonetic spelling of the word with phonetic spelling (or with approximately phonetic spelling). For instance:

Item 44. STRAT

A. straddle
B. without curves
C. avenue
D. swagger
E. historical

(In the above item, STRAT replaces "straight".)

2. Making other changes in the spelling that will not change it from phonetic to non-phonetic. For instance "k" might be substituted for "c", "j" for "g", "z" for "s", etc., where such changes are compatible with the correct pronunciation of the word.

3. Omitting vowels, where the appropriate vowel pronunciation is included in the correct pronunciation of the name of a letter in the appropriate position (e.g., "cable" might be spelled KBL. Note that the "A" sound in "cable" would be included in the pronunciation of the name of the letter "K").

4. Unstressed vowel sounds may also be omitted (e.g., "bitter" might be spelled BITR, thus omitting the unstressed vowel "E" in the second syllable).

5. In some cases even a stressed vowel (e.g., the vowel in the second syllable of "discouraged" in Sample Item 42) may be omitted, when the disguised word is sufficiently long to permit positive identification even without the vowel sound.

The additional rule was imposed that if possible the definition of the word should not be so explicit that the examinee could work backwards from the options and thus decipher the disguised word very easily. Consider, for instance, the following item:

Item 45. SIDR

A. frozen water
B. instrument for writing in ink
C. fermented grape juice
D. candy made from cacao beans
E. apple juice

(The answer is E.)

The student might have no idea what SIDR is, but he could examine the options, translate them to "ice", "pen", "wine", "chocolate", and "cider" respectively, and then easily see that Option E is the one disguised as SIDR. Therefore the following version of the item, which is perhaps a little harder, might be preferable for purposes of this test, in order to maximize the deciphering aspect and minimize the vocabulary aspect in determining the answer.

Item 46. SIDR

A. sorrowful
B. law violator
C. measurement
D. shadow
E. beverage

(The answer is E.)

It has already been pointed out that the items are not intended to measure vocabulary level directly, except to a very limited degree. Nor are they intended to measure spelling ability. (This, too, is measured directly by another test in the battery.)

What this test does measure is probably somewhat related to a kind of ability to become used to "strange" modes of spelling that would be useful in learning shorthand or a foreign language. Since the exact ways in which the spelling is distorted are not specified explicitly in the directions, what is measured, to a considerable extent, is ability to draw inferences of a type that would be drawn in learning to read a foreign language. The ability to puzzle out, from context and appearance, the meaning of a word which is vaguely reminiscent of a familiar English word is probably rather closely related to the ability measured by this test.

While scores on the test can reasonably be expected to be related to foreign language aptitude, perhaps a word of caution is in order, to the effect that this aptitude is a complex of abilities, and no single test can give a definitive answer to the question of whether someone can succeed in all aspects of learning any foreign language.

E. English Test

Purpose of test. The purpose of this test is to measure ability to express oneself adequately in English. Five separate subscores are obtained--spelling, capitalization, punctuation, English usage, and effective expression--to measure five separate aspects of ability in English.

The test is intended to provide measures of general achievement in the area of English.

Historical antecedents. Standardized spelling tests go back a long way. In 1897 several versions of a spelling test were administered to about 20,000 students (11).

Since then many different kinds of objective test items have been used to measure spelling ability. There have been some empirical studies (10) to determine whether items which test ability to recall the correct spelling of words, by requiring students to actually write out the words to show that they know how to spell them, are superior to items of the kinds used in most standardized spelling tests, which "merely" test recognition of correct spelling. These items may require the student to identify which of a long list of words are misspelled, or to decide which one of several ways of spelling a single word is the correct way, or to pick out the one word out of four or five choices that has been misspelled. The outcome of all this experimentation, in the opinion of the present writers, is that while the various types of items probably differ somewhat in their efficiency (reliability in relation to the amount of testing time required), there is little difference among them in regard to the nature of the ability measured. Neither of the two basic types of spelling test (recall and recognition) is outstandingly superior to the other. Thus the decision on which to use may have to rest on the particular administrative conditions and other circumstances (time available, scoring facilities, etc.).

There has also been considerable experimentation with different types of items to measure the other aspects of

English achievement besides spelling.

Description of Talent English Test. The Talent English Test has five separately scored parts:

1. Spelling
2. Capitalization
3. Punctuation
4. English usage
5. Effective expression

The test is primarily concerned with written English, but presumably some generalization to spoken English on the basis of the English Usage and Effective Expression subscores is justifiable.

Each of the five parts of the test is discussed in turn.

1. Spelling. Each item consists of four words, of which one may be misspelled. The task is to determine which one of the four, if any, is misspelled. Below are three sample items:

Item 47.

A. excellent
B. exercise
C. exstract
D. extreme
E. None of the above

Item 48.

A. affection
B. complexion
C. electrician
D. permission
E. None of the above

Item 49.

A. uncle
B. nickle
C. label
D. villain
E. None of the above

(The answers to these three items are C, E, and B respectively.)

The Spelling subtest is intended to measure ability to spell--not size of vocabulary. Students who do not have especially large vocabularies but who are able to spell most of the words they have encountered should get a good score. Therefore the test measures ability to spell reasonably common words--words that have probably been encountered at least once by almost all of the students. Note that the words are not necessarily ones that the students can define precisely or even loosely. (It is possible for good spellers to know how to spell words they do not exactly understand.) To insure that the vocabulary would not be too difficult, the words tested were limited to the 10,000 used most frequently in written English, according to the Thorndike-Lorge word list.* Furthermore, because this subtest is designed to measure spelling ability, not exactitude of vocabulary, no attempt is made to penalize students for misconceptions about the pronunciation of words (e.g., "larcency" for "larceny", "irrevelant" for "irrelevant", "Febuary" for "February", etc.).

---

*Thorndike, E. L., and Lorge, I., The Teacher's Word Book of 30,000 Words, Teachers College Bureau of Publications, New York, 1944.

What is actually measured directly by the test is ability to recognize correct spelling and incorrect spelling rather than the ability to recall how a word is spelled when given no visual clues (or miscues) at all.

Except in the case of such activities as proofreading and editing, the ability to spell correctly when writing is generally more important than the ability to recognize

spelling errors when one is not writing. It might appear superficially that the ability to recognize spelling errors is all that is measured by this test. It is believed, however, that the ability to recall correct spelling--in other words actually to spell correctly--is also measured, since the correlation between the two is undoubtedly high, and since the test was administered in such a way that the correlation is probably enhanced. (This was achieved by providing the students with scratch paper--for the convenience of those students who, when they are in doubt about the correct spelling of a word, like to write it out in the various ways under consideration.)

Thus what is measured directly by the test is ability to recognize spelling errors when presented with the visual stimulus, while the actual ability to spell a word correctly when writing it is probably tapped indirectly. No effort is made to measure the ability to spell orally in the absence of any visual stimulus. This ability is inherently a rather unimportant aspect of spelling. Furthermore, it seems probable that it has a sizable correlation with the spelling abilities that are measured.

In the interests of maximizing the efficiency of the battery as a whole, an effort was made to keep to a minimum the extent to which spelling scores would be correlated with general intelligence, which is covered adequately by other tests in the battery. In order to keep the correlation with intelligence at a minimum, phonetic spelling errors were used (i.e., errors which would still permit the word to be pronounced correctly if it were pronounced as spelled) rather than non-phonetic spelling errors.

Elimination of non-phonetic errors is justified on the grounds that those who make such errors are also prone to phonetic errors if these are the only kind offered. This is in accordance with the findings of Dailey and Mays. (5)

(An example of a phonetic error is "exstract" for "extract". A non-phonetic error would be "ectract" for "extract".) When words which are not pronounced as they are spelled were included, the misspelling was in the direction of making the word more phonetic (e.g., atturney, sargeant).

The emphasis in all cases was on words that present a problem of some sort in regard to their spelling, rather than on words whose spelling is perfectly straightforward and present no difficulties.

A spelling test was included in the battery for two reasons. In the first place, the kind of achievement measured by this test is inherently important. The ability to spell adequately is necessary in order to be able to write acceptable English. This ability is essential for many types of jobs--e.g., secretary, typist, etc. The ability to recognize spelling errors is also directly necessary in many job activities--e.g., proofreading, editing, etc. Most fundamentally, however, correct spelling is a central part of the manifest culture of an educated person.

The second major reason for including spelling in the battery was that its correlation with many other aptitude

tests is quite low, and that therefore it could be expected to contribute a large unique component. Relatively little is known about what spelling ability actually consists of (e.g., to what extent it depends on visual memory, visual perception, auditory-visual association, etc.), but there <u>is</u> evidence that its inclusion in a battery contributes significantly to the battery's validity for a number of purposes.

2. <u>Capitalization</u>. This test consists of a passage printed entirely in small letters. The task is to determine which of the words with numbers under them should start with a capital letter.

Below is a sample passage with 36 items:

```
george hathaway spent last november in paris
50     51             52   53          54

with his brother henry and their young cousin,
     55  56      57        58    59    60

louis beauregard; they stayed in rather a large
61    62          63               64     65 66

hotel overlooking the seine river.  in december
67    68          69  70    71      72 73

they all went to england (to london, i believe),
     74          75       76 77      78

where they spent christmas with dr. and mrs.
79    80   81    82              83      84

hathaway.
85
```

(Answers: The words to be capitalized are Nos. 50, 51, 53, 54, 57, 61, 62, 70, 71, 72, 73, 75, 77, 78, 82, 83, 84, 85.)

3. <u>Punctuation</u>. The Punctuation subtest has two parts--Section <u>a</u> (Punctuation Marks) and Section <u>b</u> (Sentence Structure).

In Section <u>a</u>, each item consists of a sentence in

which a section is printed without punctuation. This section of the sentence is then presented punctuated in several alternative ways. The task is to determine which of the ways presented is correct. A sample item follows:

Item 86. How blue the sky has turned

A. sky has turned.
B. sky has turned!
C. sky has turned?
D. sky, has turned.
E. sky; has turned.

(The answer is B.)

The purpose of Section a is to test knowledge of the appropriate use of all standard punctuation marks: period, comma, semi-colon, colon, apostrophe, single quotes, double quotes, question mark, exclamation point, dash, parentheses, hyphen. While the examinee is required to differentiate between correct and incorrect punctuation, he is not required to determine which of two or more acceptable or defensible manners of punctuation is preferable. For instance, no effort is made to test whether the student recognizes cases where although a period is grammatically correct, a semi-colon would be preferable; nor is any effort made to test whether the student has the same preference as the item writer in cases where either commas, dashes, or parentheses are technically acceptable. Such decisions involve refinements of style, and as such are matters about which there might be considerable difference of opinion among expert writers. Knowledge of correct capitalization is not tested in these items, since it is covered in a separate section of the test.

In Section b, each item consists of either a fragment of a sentence, a complete sentence, or two or more sentences

run together. The task is to determine which it is.

The directions are to mark the answer sheet as follows:

Mark "0" if it is part of a sentence.

Mark "1" if it is one complete sentence.

Mark "2" if it is two or more sentences run together.

Four sample items follow:

Item 87. A course of study including English, French, algebra, American history, and chemistry.

Item 88. The football team having tied one game and won all the rest.

Item 89. Try this, it will help you.

Item 90. Do you know how?

(The answers are 0, 0, 2, and 1 respectively.)

The function of this section is to determine whether the student has mastered the concept of what constitutes a sentence--including the fact that every sentence or clause has a subject and a predicate. Mastery of this concept is essential in order to be able to write acceptable English.

4. English Usage. Each item consists of a sentence with a word or group of words missing, the missing section being represented by a blank. The task is to determine which of several (three to five) choices fits in the blank best. The instructions specify that if two choices are correct, the better one is to be selected.

Here are two sample items:

Item 91. Bob __________ arrange it.

A. might could
B. might be able to
C. could maybe
D. maybe could
E. would maybe be able to

(The answer is B.)

Item 92. Ed and ________ planning to go.

A. myself was
B. me was
C. I was
D. myself were
E. I were

(The answer is E.)

The items measure knowledge of correct English usage and knowledge of the best way of expressing a given idea. Knowledge of vocabulary is not to be measured in this section, since that is measured elsewhere in the battery. What is to be measured is knowledge of preferred usage. The student is not required to determine in absolute terms whether any particular usage is definitely incorrect or whether on the other hand it is marginally acceptable and merely undesirable. Instead, the student is required to decide which of several options provides the best (grammatically correct, clearest, most unambiguous, least awkward) way of expressing a particular idea. In Sample Item 91, for instance, although Options C, D, and E are not technically incorrect, they do represent an unidiomatic and therefore awkward word order. ("Maybe Bob could arrange it" would be a less awkward word order than any of the sequences in which the word "Bob" precedes the word "maybe".)

5. Effective Expression. Each item consists of three, four, or five sentences, each expressing the same idea, but only one of them expressing it well. The student is to pick which of the sentences expresses the idea best. (He is given no hint as to what criteria to use in making his decision.)

A sample item follows:

Item 93.

A. Things such as this helped our forefathers push back the wilderness, and create the richest nation in the world.

B. Our forefathers created the richest nation in the world, helped by things like this to push back the wilderness.

C. Things such as this helped our forefathers in their pushing back of the wilderness and in their creation of the richest nation in the world.

D. To create the richest nation in the world, our forefathers, being helped by things like this, pushed back the wilderness.

(The answer is A.)

These items are used to measure recognition of good prose expression--i.e., sentences which express ideas clearly, concisely, and smoothly. To do well on this subtest, the student must reject sentences which express an idea ambiguously, awkwardly, or unduly wordily. He is not required to determine whether the preferred sentence represents the best possible way of expressing the idea. He need only choose among the sentences given.

The items do not test the ability to recognize simple grammatical errors, since this is covered in the English Usage subtest, nor do they test knowledge of punctuation, since this likewise is covered in a separate score.

E. Word Functions in Sentences Test

Purpose of test. This test is intended to measure a person's sensitivity to grammatical structure, whether or not he has had formal instruction in the rules of grammar.

Historical antecedents. This type of test was originated only a few years ago, by John B. Carroll, who designed

it as a measure of foreign language aptitude, and incorporated it in his Psi-Lambda Foreign Language Aptitude Battery as Test 4 (Words in Sentences). It has proved to have excellent validity as a predictor of foreign language learning ability. For instance, using the same criterion (course grades in an intensive Air Force "trial course" in Mandarin Chinese) as has already been mentioned in connection with the Paired Associates Test of the Psi-Lambda Battery, Carroll demonstrated that the Words in Sentences Test helped predict this criterion, and attributed this result chiefly to a common "linguistic interest" factor.

<u>Description of Talent Word Functions in Sentences Test</u>. This test uses the same type of item as the Words in Sentences Test (Test 4 of the Carroll-Sapon Psi-Lambda Battery), mentioned above.

Each item consists of a "key sentence", followed by one or more other sentences. The key sentence always has a word or phrase printed in capital letters, and the task is to determine which of the five underlined words or phrases in the following sentence(s) performs the same function in its sentence as the word or phrase in capitals performs in <u>its</u> sentence.

Three sample items follow:

<u>Item 94</u>. They walked GAILY down the street.

The <u>newly</u> (A) arrived couple <u>used</u> (B) up their <u>money</u> (C) <u>too</u> (D) <u>fast</u> (E).

Item 95. HAVING an optimistic nature, she expected things to turn out well.

Driving a car requires a license.
A

The weather having finally cleared up,
B

Mrs. Rogers was able to work in

her garden.

Helen, busily planning her afternoon,
C

wasn't watching where she was going.
D E

Item 96. This is the WAY to do it.

I do it a different way.
A

I use a different method.
B

The way to do it is like this.
C

That's the wrong method.
D

There are many ways of doing it.
E

(The answers to these three items are E, C, and D respectively.)

This is a very difficult test for most students. It was made difficult deliberately, since there is reason to believe a hard test would predict ability to learn foreign languages better than an easier one.

Although formal training in grammar and previous study of foreign languages (particularly of highly inflected ones, like Latin) are probably of some help on this test, the fact that the terminology of grammar is not used at all contributes to reducing the effects of formal training to a minimum. To perform well on the test, one must understand

something about the structure of a sentence, and recognize the function of each word or phrase in the sentence.

The ability measured by the test is probably related to the ability to learn the grammar and characteristic sentence structure of a foreign language--particularly a heavily inflected language (such as Latin, for instance) in which word order is typically very different from that of English. It is probably also related to the ability to learn the formal rules of English grammar.

A relationship has been demonstrated between scores on this type of test and the ability to learn a foreign language. Nevertheless, since this test measures only one aspect of that very complex ability, a high score on the test does not guarantee success in mastering a foreign language. But considered in conjunction with other relevant tests, the Word Functions in Sentences Test should prove very useful as a predictor.

G. Reading Comprehension

Purpose of test. The purpose of the test is to measure ability to comprehend written materials dealing with various kinds of subject matter.

Historical antecedents. Although numerous types of standardized reading tests have been used over the last 40 years, by far the most commonly used type is the one that consists of passages to be read, with each passage followed by several multiple-choice questions testing comprehension. In answering the questions the student is usually permitted to refer back to the passage as often as he likes.

Description of Talent Reading Comprehension Test. The Talent Test is of the type described in the preceding paragraph. The passages cover a wide range of topics, including some dealing with social studies, some dealing with natural science, and some which are of a literary character. Poetry as well as prose is included.

A poem and some sample items based on it follow:

(1) Loveliest of trees, the cherry now
(2) Is hung with bloom along the bough
(3) And stands about the woodland ride
(4) Wearing white for Eastertide.

(5) Now, of my threescore years and ten,
(6) Twenty will not come again,
(7) And take from seventy springs a score,
(8) It only leaves me fifty more.

(9) And since to look at things in bloom
(10) Fifty springs are little room,
(11) About the woodlands I will go
(12) To see the cherry hung with snow.

Item 97. What kind of tree is the poem about?

A. White pine
B. White birch
C. White oak
D. Cherry
E. None of these

Item 98. What does the word "ride" probably mean, as used in line 3?

A. Auto ride
B. Tree
C. Horseback ride
D. Flowers
E. Road

Item 99. What season is it, in the poem?

A. Winter
B. Spring
C. Summer
D. Fall
E. There is no way of telling.

Item 100. The poet indicates, a little ruefully, that his life

A. will pass too quickly.
B. will probably end very soon.
C. has been unhappy.
D. has been devoid of beauty.
E. has been uneventful.

Item 101. How old does the poet say he is?

A. 20
B. 30
C. 40
D. 50
E. 70

Item 102. When the poet says he "will go" (line 11), he probably means he will go

A. on foot.
B. by boat.
C. by automobile.
D. by horse-drawn sleigh.
E. on snowshoes or skis.

Item 103. The poet expresses a liking for looking at

A. snow-covered trees.
B. snow-covered ground.
C. spring blossoms.
D. autumn leaves.
E. Easter lilies.

Item 104. What does the poet say are the loveliest trees?

A. Any white trees
B. Any trees in bloom
C. Cherry trees
D. Any woodland trees
E. He doesn't say.

(The answers are D, E, B, A, A, A, C, and C respectively.)

It is important to note that what the test measures is the ability to read with comprehension, rather than mere ability to mouth or recognize the printed word without understanding the fact, idea, or concept that the writer is attempting to convey.

Reading comprehension skills have been classified into the following nine categories:*

---

*Davis, F. B., "Fundamental Factors of Comprehension in Reading", Psychometrika, 1944, 9, 185-197.

1. Knowledge of word meanings.
2. Ability to select the appropriate meaning for a word or a phrase in the light of its particular contextual setting.
3. Ability to follow the organization of a passage and to identify antecedents and references in it.
4. Ability to select the main thought of a passage.
5. Ability to answer questions that are specifically answered in a passage.
6. Ability to answer questions that are answered in a passage but not in the words in which the question is asked.
7. Ability to draw inferences from a passage about its contents.
8. Ability to recognize the literary devices used in a passage to determine its tone and mood.
9. Ability to determine a writer's purpose, intent, and point of view, i.e., to draw inferences about a writer.

Of these nine categories, all but the first is covered directly by items in this test. Knowledge of word meanings is measured in the Information Test Vocabulary Scale and therefore is not measured directly in the Reading Comprehension Test. To a certain extent, however, it is unavoidable that it will affect Reading Comprehension scores, since the vocabulary of the passages and questions must be reasonably well understood if the items are to be answered correctly. But extremely difficult vocabulary has been systematically avoided.

In general, the vocabulary is within the first 15,000 words on a word frequency count such as the Thorndike-Lorge list.* The primary exceptions to this limitation would be

---

*Thorndike, E. L., and Lorge, I., The Teacher's Word Book of 30,000 Words, Teachers College Bureau of Publication, New York, 1944.

words whose meaning is explained in the passage and words whose meaning can be inferred reasonably well from the context.

None of the items is answerable without reading the passage. For example, items to measure whether the student understands the meaning of a word used in the passage are not included unless the distractors would be justifiable if the words were used in a different context. (See Sample Item 98 for an example of this.)

The ability measured by this test is closely related to general intelligence and as such is a good predictor of school success in an academic or liberal arts curriculum.

In the average secondary school situation, most students who are poor readers probably lack the basic intellectual potential to become good readers. This generalization may [illegible] somewhat less applicable in inadequate school situations. In the unusual situation where a student scores low on the Reading Comprehension Test but obtains good scores on other types of intellectual tests, he may have a specific reading disability that will respond well to special remedial teaching.

H. Creativity

Purpose of test. The purpose of the test is to measure the ability to find ingenious solutions to a variety of practical problems. It is designed to measure something over and above tests of logical reasoning where the problems are very highly structured and have formal unique solutions. There is some evidence that this ability is related to many kinds of creativity.

Historical antecedents. Many attempts have been made

to measure creativity or ingenuity. Perhaps the most comprehensive effort in recent years has been that of Guilford (9) who studied the interrelations of many tests of fluency, imagination, and originality. A recent development is the new item type designed by Flanagan (7) to measure ingenuity and inventive skill, the ability to devise ingenious procedures, equipment, or presentations. He developed a new response format which avoided suggesting the correct responses to the problem as presented. The examinee is given the problem and a few clues and is required to devise an ingenious solution. This is the type of item used in the Ingenuity Test of the Flanagan Aptitude Classification Tests (FACT Battery).

Description of Talent Creativity Test. This test is modeled on the FACT Ingenuity Test described above. Each item consists of a complex problem similar to a practical problem that might be encountered in real life. The examinee must think of a clever or ingenious solution. The correct solution (word or phrase) seems to "click" into place, and the examinee has a definite feeling of closure when he identifies the solution. The choices are given in terms of the first and last letters of possible right answers. This is to insure that the examinee really develops the solution, and doesn't just pick it out from among five alternatives. Two sample items are given below:

Item 105a. Attic and cellar stairways and stepladders are often causes of household accidents. One method for preventing slipping and falling on wooden stairs or steps is to apply a coat of varnish to each step, and while the varnish is still slightly sticky, apply a small amount of

A. d - - n to each step.
B. s - - d to each step.
C. r - - g to each step.
D. g - - n to each step.
E. r - - t to each step.

(The answer is B, since "sand" is the solution.)

Item 105b. The sanding and smoothing of knobs and other small round wooden objects is always a problem because flat sheets of sandpaper do not fit the knobs. The sandpaper buckles and tends to wrinkle. One solution is to cut

~~several~~

A. e - - - - - - - n d - - - s in the sandpaper.
B. h - - - - - - - l t - - - s in the sandpaper.
C. p - - - - - - - l s - - - s in the sandpaper.
D. t - - - - - - - n w - - - s in the sandpaper.
E. m - - - - - - - h b - - - s in the sandpaper.

(The answer is C, since "parallel slits" is the solution.)

The items are not intended to require detailed knowledge of specialized fields. They <u>are</u> intended to provide a measure of something beyond general knowledge, vocabulary, and deductive reasoning ability. And there is evidence that they do. As would be expected from an examination of the items, the Creativity test to some extent does measure general ability of the type measured by intelligence tests. And it has been shown to have a moderate amount of overlap with various kinds of reasoning tests, including mechanical reasoning. However, it does have an appreciable amount of "unique reliable variance"; in other words part of what it measures is an ability not covered by any other kinds of verbal or reasoning tests that have been tried out in conjunction with it.

High scores on this test should be interpreted as indicating inventiveness or creative ingenuity. Such scores would indicate that the individual should do well in dynamic problem-solving activities involving the creation of original solutions to challenging, complex problems.

I. Mechanical Reasoning

Purpose of test. The purpose of this test is to measure ability to visualize commonly encountered objects and to recognize how the operation of everyday physical forces and principles (e.g., gravitation, pressure, equilibrium) and basic kinds of mechanisms (e.g., gears, pulleys, wheels, springs, levers) would affect them. Tests like this have sometimes been called measures of "barnyard physics". A kind of reasoning which is related to mechanical aptitude is involved in a test of this sort.

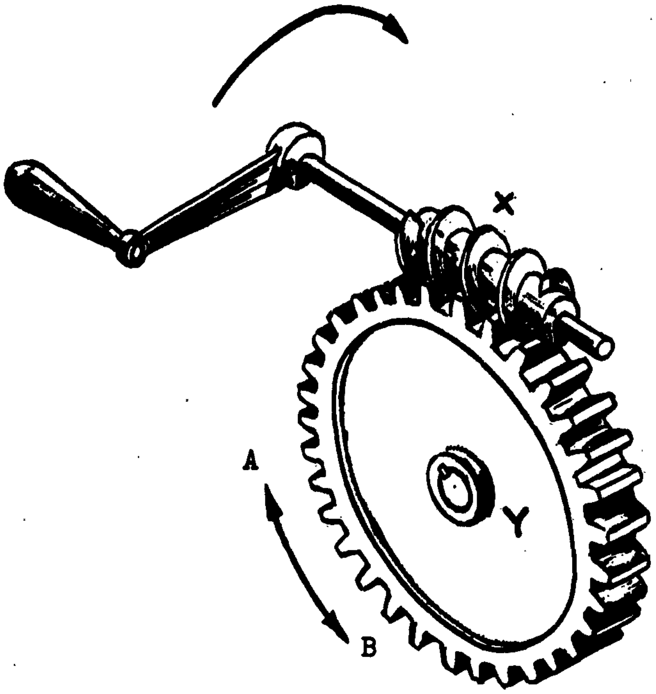
Historical antecedents. As early as World War I, mechanical ability was recognized as a type of testable ability separate from general intelligence. Early versions of mechanical ability tests were constructed at that time. A little later, various combinations of manipulative, pictorial, and verbal items were labelled as mechanical ability tests by a number of investigators. It was not until much later that it was recognized that the different types of "mechanical" items were really measuring somewhat different mental factors. In 1940 the Bennett Mechanical Comprehension Test (1) was developed. This was the first mechanical reasoning test of the pictorial type.

By the end of World War II the broad area of mechanical ability had been resolved into such commonly used tests as mechanical information, electrical information, tool usage, spatial visualization, eye-hand coordination, and pictorial mechanical reasoning. The latter, in the form employed by Bennett, has been the most widely used. However, many

comprehensive aptitude test batteries employ most or all of these types of tests. Despite the relatively high correlation between scores on them, all have proven to yield some unique contribution toward predicting success in mechanical occupations. The kind of ability measured by a mechanical reasoning test of the pictorial type is only one of the abilities requisite for success in certain mechanical occupations. This must be recognized in interpreting results. An engineer, for instance, may also need a high degree of verbal and mathematical ability.

There is a wealth of evidence supporting the validity of mechanical reasoning tests (4) for a wide variety of occupations, such as airplane pilot, auto mechanic, machinist, welder, aircraft mechanic, sheet metal worker, tool and die maker. Such tests have been widely used by the Armed Forces, both for initial selection and for assignment to training. In the very extensive validation studies in the Armed Forces, tests of this sort have almost invariably been found to correlate substantially with measures of success in mechanical training courses. These results are confirmed by many similar studies in industry.

Description of Talent Mechanical Reasoning Test. This test uses about the same type of items as were used in Bennett's Mechanical Comprehension Tests. Each item consists of a drawing, with a multiple choice question concerning how the objects pictured behave. A sample item is shown below.



Item 106. When worm gear X is turned in the direction shown, in which direction, if any, does gear Y turn?

A. In direction A
B. In direction B
C. First in one direction and then in the other
D. It doesn't turn

(The answer is A.)

In developing the items an effort was made to make them answerable by persons who had not had any specific formal training in physics, or esperience in working with motors, woodworking, or other crafts.

Knowledge of tools and knowledge of electricity are not measured.

The test measures the ability to determine, through visualization and reasoning ability and on the basis of past observations of common mechanical forces at work, how physical objects behave. While a course in physics

may help on many of the items, a youngster with no formal training in physics can answer the questions intuitively or through reasoning if he has a bent towards activities of a mechanical nature. Nevertheless, in interpreting scores on this type of test, past training and experience should be considered. The fact that girls, on the average, score considerably lower than boys is consistent with the hypothesis suggested above that environmental factors (training, experience, etc.) play a significant role in the development of the kind of ability measured by this test. However, this circumstance does not render the test any less useful.

13a. Visualization in Two Dimensions
and
13b. Visualization in Three Dimensions

Purpose of Tests. The general purpose of these two tests is to measure spatial visualization.

Historical antecedents. Spatial visualization has been measured by a great many different kinds of tests--tests involving manipulation (diagrammatically) of flat figures in two dimensional space, for instance rotation of the figures, and tests involving manipulation (diagrammatically) of three-dimensional figures in three-dimensional space, for instance folding of flat figures to make three-dimensional figures, unfolding of three-dimensional figures to flatten them out (sometimes called surface development), rotation of solid figures such as cubes, visualization of hidden parts of solid structures, (as in a "block-counting test" where the task is to determine how

many cubical blocks are in an irregularly shaped structure not all of which is visible).

While the tasks called for by these various types of items differ widely, it has been found, by means of factor analysis research, that most of them measure essentially the same kind of ability. (8)

This ability has been found to be involved in many occupational specialties of a technical, mechanical, or engineering nature--for instance, design engineer, carpenter, mechanical draftsman, airplane pilot. Tests of this sort are useful for selecting students who are likely to do well in shop courses. These tests are also useful in vocational situations where the individual must make a quick adaptation to the demands of certain semi-skilled jobs - especially those involving the assembly of irregularly shaped parts, machine operation, and the packing of objects of different shapes and sizes.

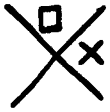
Description of the Talent Visualization Tests. It was decided to measure spatial ability with two separate tests in the Talent battery, a two-dimensional visualization test and a three-dimensional one, although it was expected that the resultant scores would turn out to be rather highly correlated. Each of these two tests will be discussed in turn, after one feature they have in common is mentioned. This is the fact that the diagrams were all made quite large, and free of fine detail, so that exceptional visual acuity would not be required to determine the answer.

1. Visualization in Two Dimensions. This test is of the type that calls for distinguishing between figures

which have been "reflected" and figures which have merely been rotated without reflection, so that they are essentially identical, merely in different positions. Each item consists of a given figure followed by five other figures. The task is to determine which one of the five could be twisted around to coincide with the given figure, without "lifting it". (The other four are "mirror images" and would have to be reflected as well as rotated. A sample item follows:

Item 107.

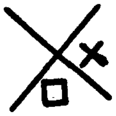
A



B

C

D

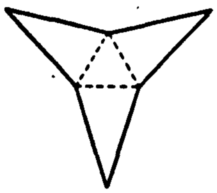


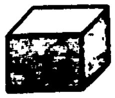
E

(The answer is B.)

2. Visualization in Three Dimensions. This test measures ability to visualize what a two-dimensional figure would look like if it were folded or rolled to form a three-dimensional figure. Inthese items folds are indicated by dotted lines and cuts by interior solid lines. Two sample items, an easy one and a relatively hard one, are shown below.

Item 108.

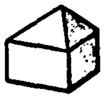




A

B

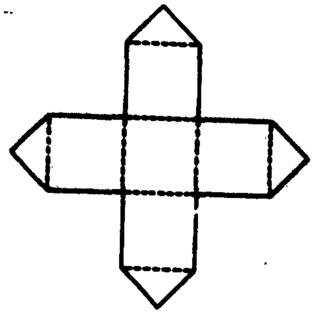
C

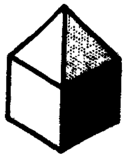


D

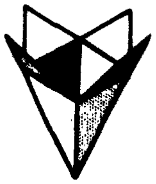
E

Item 109.



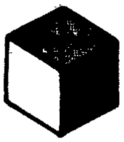


A

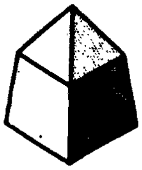


B

C



D



E

(The answers are E and D, respectively.)

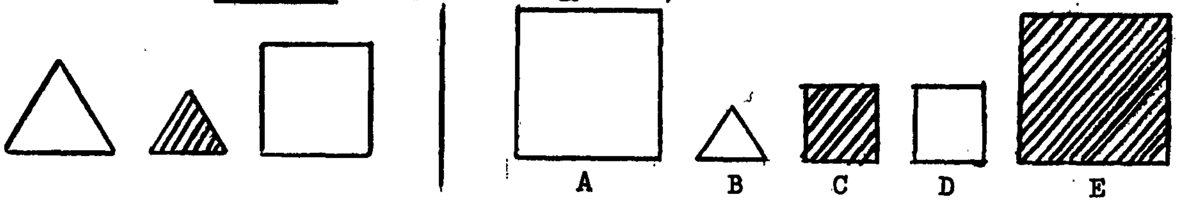
14. Abstract Reasoning

Purpose of Test. This is a non-verbal test of inductive reasoning ability.

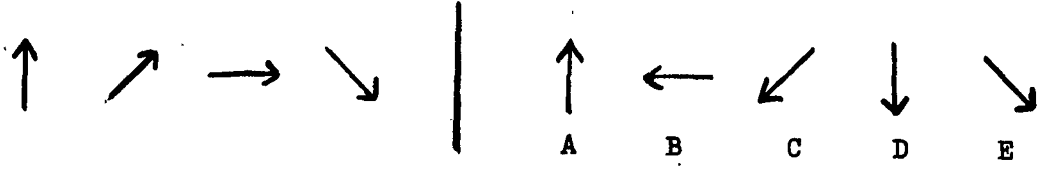
Historical antecedents. One of the major purposes of non-verbal reasoning tests originally was to measure the basic intelligence of individuals with language handicaps. Efforts along these lines were made even before World War I. It has since been found, however, that the kind of abstract reasoning ability measured by non-verbal tests is not the same thing as the kind of ability measured by verbal intelligence tests. It is this very fact that makes this kind of test more generally useful than it would be if its applicability were limited to persons with language handicaps. Many currently used intelligence tests include measures of abstract non-verbal reasoning, as do most multiple aptitude batteries. There are several different types of items measuring ability to do abstract reasoning about diagrammatic material that have been widely used. Among these types are figure analogies items, figure sequences items, pattern matrices items, and figure grouping items. The first three of these item types all consist of a pattern of diagrams, with one missing portion to be selected from among several options. In the fourth item type, figure grouping, the task is a little different; it is to determine what the principle is which ties several diagrams together, and on the basis of this principle to determine which pattern from among several options belongs with the group of patterns. In all of these abstract reasoning item types, the solution of the

item depends upon the ability to determine a logical relationship among elements of the pattern and to apply this relationship in order to identify an element that belongs in the pattern--either in a specified position, as in the case of the figure analogies, figure sequences, and pattern matrices tests, or, as in the case of the figure grouping test, just some place in the group. Sample items of each type are shown below.

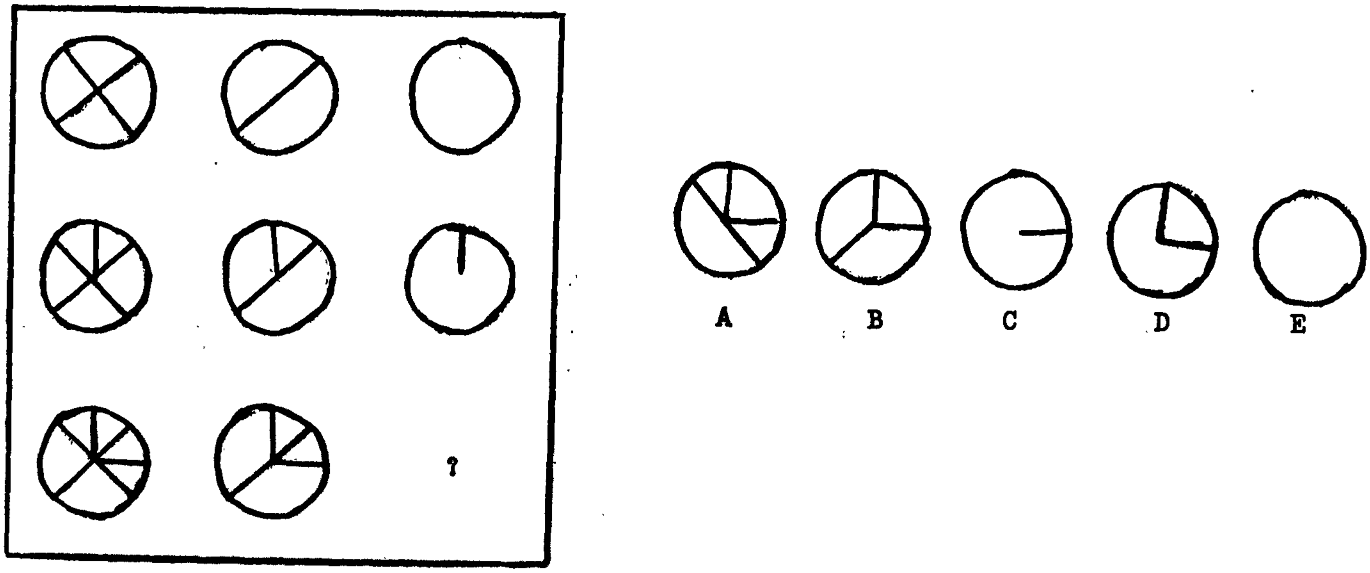
Item 110 (Figure analogy item).



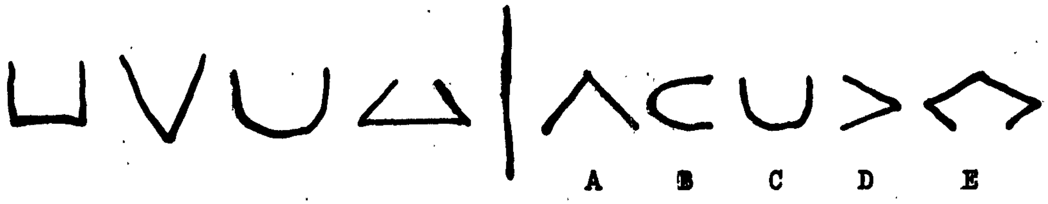
Item 111 (Figure squence item).



Item 112 (Pattern matrix item).



Item 113.

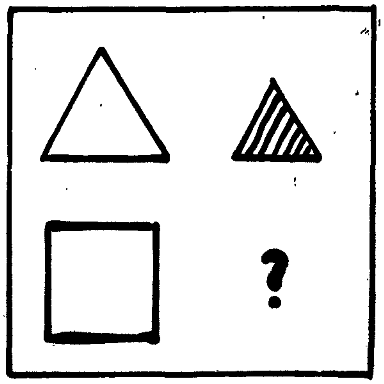


(The answers are C, D, D, and C respectively)

On the basis of empirical data there seems to be relatively little clear evidence that any of the types of items discussed above is markedly superior to the others.

Description of Talent Abstract Reasoning Test. It was decided to use the pattern matrices form of item in the Talent test because it seemed to have the advantage of somewhat greater generality than the other item types. Essentially both figure analogies and figure sequences are special cases of pattern matrices. Almost any figure analogy item can be translated into the pattern matrices format, using a 2 x 2 matrix. For instance, figure analogy item 110, in the pattern matrices format, would appear as shown in sample item 114.

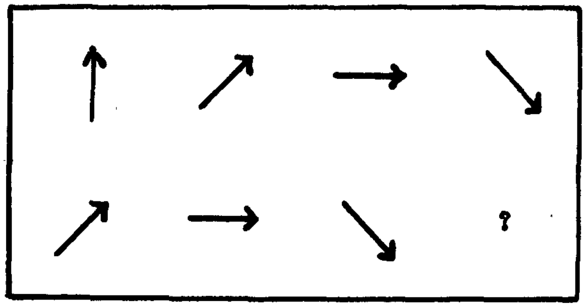
Item 114.



Same options as
Item 110.

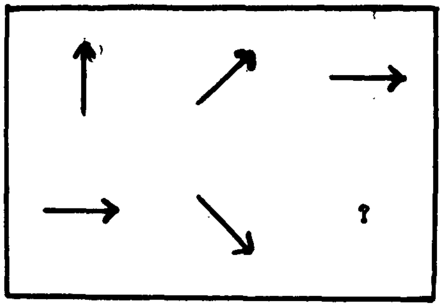
Similarly, the figure sequence item type is readily convertible to the pattern matrices format. For instance, consider figure sequence item 111. Two pattern matrices versions of the same item are shown, as sample items 115 and 116.

Item 115.



Same options as Item 111.

Item 116.



Same options as Item 112.

(The answer is D for both items above as for Item 111.)

Solution of abstract reasoning items depends on ability to determine inductively the logical relationship among the elements of the pattern, and to apply this relationship in order to identify the missing element.

While no claim can be supported that a test of this type is wholly culture-free it is certainly more "culture-fair" than a test directly involving verbal ability would probably be.

Scores on this type of test are widely used as a measure of the ability to do good school work in academic curricula, especially in the case of individuals with little schooling or with cultural or environmental handicaps. Individuals with these handicaps who score low in vocabulary and reading but high on this test can be regarded as having a high innate potential for academic work. However, if plans are being made for advanced academic work low scores on the other tests cannot be disregarded, since they indicate that considerable remedial work should precede advanced schooling. Among the general run of students, scores on the Abstract Reasoning Test should be interpreted as just another element in the general domain of intellectual potential.

L. Arithmetic Reasoning (Mathematics Part I)

Purpose of Test. This test is designed to measure the ability to do the kind of reasoning required to solve arithmetic problems. Insofar as possible, arithmetic computation, except at the very simplest level, is excluded by the design of the test.

Historical Antecedents. In addition to their obvious use for achievement tests in elementary school, where arithmetic reasoning is taught, objective arithmetic reasoning items have frequently been used in intelligence tests.

Description of Talent Arithmetic Reasoning Test. The chief difference between the Talent test in arithmetic reasoning and most other tests of this ability is that in the Talent test a systematic effort was made to eliminate almost all computation, even at an easy level. This was done in an effort to reduce the correlation between scores on this test and scores on the Arithmetic Computation test--a correlation which is usually quite substantial. The aim of the Talent test is solely to determine whether the student can do the reasoning part of a problem--not whether he will do the subsequent computation carefully enough to avoid careless errors, nor whether he knows the simple number combinations (e.g., 9 x 7 = 63). The intention was to measure these latter facts and abilities in the Arithmetic Computation test and to try to avoid contaminating the Arithmetic Reasoning test with them. Several devices were used to achieve this end. One such device was to require the student to tell how he would solve a stated

problem but not to require him to actually do the necessary computation. A sample item is shown below:

> <u>Item 117</u>. Three percent of a certain number is 141. To find the number, you would
>
> A. divide 141 by .03
> B. divide 141 by 3
> C. divide 141 by 300
> D. multiply 141 by .03
> E. multiply 141 by 300
>
> (The answer is A.)

A partial list of the types of items used follows:

1. Items asking what process (e.g., addition, subtraction, multiplication, or division) should be used for a given problem. (See sample item 117.)

2. Items which present several facts and then ask which one is <u>not</u> needed in sovling a certain problem.

3. Items which present several facts and then ask what additional information would be needed to solve a certain problem.

4. Items which present several facts and then ask which one of several problems cannot be solved on the basis of the facts given.

No items were included which required algebraic techniques for solution. However, no problem was proscribed just because it <u>could</u> be solved efficiently by algebra, provided there was an alternate way of solving it, without algebra.

Arithmetic reasoning is perhaps the sort of test in which the relationship between aptitude measures and achievement measures, which sometimes appears to be a confused one, can best be clarified.

An arithmetic reasoning test measures achievement in something which has been taught, since virtually everyone in our culture today goes to school long enough to be

exposed to arithmetic problems at the sixth and seventh grade level. But it also measures aptitude--the aptitude for more advanced mathematics and the aptitude for many kinds of jobs. The ability to handle arithmetic reasoning problems well not only is directly important in many jobs but it has been known to be indicative of ability to succeed in many other kinds of jobs, that do not involve arithmetic reasoning tasks and have no apparent direct connection with mathematics.

16a. Mathematics Part II (Introductory)
and
16b. Mathematics Part III (Advanced)

Purpose of Test. The purpose of these two subtests is to measure achievement in all kinds of mathematics taught in high school (or earlier) with the exception of the areas covered in the Arithmetic Computation test and in Mathematics Part I (Arithmetic Reasoning).

Part II (Introductory Mathematics) covers topics normally taught in Grade 9 or earlier. Part II (Advanced Mathematics) covers topics normally taught in Grades 10-12 in college preparatory courses.

Historical Antecedents. Test in college preparatory mathematics, such as algebra, geometry, or trignometry, are in most situations more appropriately regarded as measures of achievement than of aptitude, since amount of exposure to this type of subject matter is highly variable, in contrast with arithmetic reasoning problems, to which virtually everyone is exposed in elementary school. Within groups where more advanced mathematics is taken by everyone, for instance in schools where it is a required

subject, mathematics achievement measures may be used to indicate basic potential for that sort of mental process.

But advanced mathematics in high school is far more likely to be an elective than a required subject. Thus these two subtests are to be regarded primarily as achievement tests. Although there is little precedent for including such tests in what may be regarded as primarily a predictive battery rather than an achievement battery, they should serve a useful purpose, particularly since achievement in mathematics has been found to be quite closely related to success in many kinds of college curricula.

Description of Talent test. The items are intended primarily to test understanding of basic concepts and methods, rather than rote memory. Little emphasis is placed on facts and definitions since this domain is covered in the Mathematics scale of the Information Test.

In interpreting results it should be borne in mind that the distinction between topics taught in Grade 9 and earlier and those taught later than Grade 9 is not a clear one since curricula in different schools differ.

1. Part II (Introductory)

   The primary emphasis of this subtest is on ninth grade elementary algebra, but other topics include some taught in elementary and junior high school, such as fractions, percentage, intuitive geometry, elementary measurement formulas (area of triangle, circumference of circle, etc.), square root.

   Actual computation is kept to an absolute minimum.

Two sample items follow.

Item 118. The area of a rectangle s inches wide is three times the area of a square with a side s inches long. The length of the rectangle is

A. $\frac{s^2}{3}$

B. $3s$

C. $3s^2$

D. $\frac{s}{3}$

E. $s\sqrt{s}$

Item 119. What is the first digit of the square root of 485726? (Hint: Do not compute the square root.)

A. 1
B. 2
C. 4
D. 6
E. 7

(The answers are B and D respectively.)

2. Part III (Advanced)

Though these items may be regarded as "achievement" items, they are not designed to measure achievement in any specific course. Therefore an effort was made to limit the items in this subtest to ones of such a nature that if considerable time had elapsed since the student studied the material, and if he had consequently forgotten it, he would be able to reconstruct it fairly rapidly, through logical analysis, provided he really understood it thoroughly in the first place. Thus for instance in measuring achievement in advanced algebra, items on how to solve a cubic or quartic equation would certainly not be included. Remembering these procedures would require

memorization of complicated formulas, and if these had been forgotten they could not be reconstructed instantaneously, even by someone good at mathematics. On the other hand, items requiring the student to use the principle that the probability of two independent events is the product of their individual probabilities would be acceptable. Even if the student hasn't memorized the theorem, it should be quite obvious to him if he understands the concept of probability.

The items sample a wide range of concepts covered in the last three years of high school mathematics. The concepts are not limited to those taught in <u>all</u> high schools offering four years of mathematics. Among the areas included are plane geometry, solid geometry, algebra (quadratic equations and beyond), trigonometry, elements of analytic geometry, and introductory calculus. Sample items follow.

<u>Item 120</u>. If $10^m = r^4$, the value of $\log_{10} r$ is

A. $4m$
B. m^4
C. 4^m
D. $\frac{m}{4}$
E. $\sqrt[4]{m}$

<u>Item 121</u>. If $a < b < c < e$ and $d > b$, it follows that

A. $d < c$
B. $a > d$
C. $d > a$
D. $c < d$
E. $d < e$

Item 122. A ratio which is always the same is the ratio of a circle's

A. area to its circumference.
B. radius to its area.
C. circumference to π.
D. diameter to π.
E. radius to its circumference.

(The answers are D, C, and E respectively.)

This subtest is intended primarily to provide a national inventory of students who have reached specified levels of achievement in mathematics.

As with the other tests, norms were obtained for each grade separately. Thus there is a suitable base for interpretation of 9th grade results, taking into consideration the fact that most of the 9th graders will not have had formal instruction on the topics covered.

17. Arithmetic Computation

Purpose of Test. The purpose of the test is to measure speed and accuracy of basic computation operations (addition, subtraction, multiplication, and division) with whole numbers.

Historical antecedents. Arithmetic computation tests can vary in several dimensions, including:

1. Whether the operations are limited to whole numbers or whether fractions and decimals are included.
2. Whether the problems are limited to those that can readily be done mentally (e.g., operations with one-digit numbers) or whether they are more difficult than that.

3. Whether the kinds of operations are limited to the four basic ones or whether more difficult operations, such as square roots, are included.
4. The generosity of the time allowance.

When arithmetic computation tests are intended primarily as school achievement measures, the chief concern is usually accuracy and knowledge of methods, rather than speed. For such purposes, therefore, it is advantageous for the test to cover a comparatively broad variety of operations, to include some problems that are innately quite difficult, and to have a liberal time allowance.

It has been found, however, that when an arithmetic computation test is primarily intended to serve as a job aptitude measure, the most useful results are obtained if the number of items is large and the time limit so short that nobody can finish, and if the items themselves are innately quite easy, so that if errors are made they are likely to be due to carelessness rather than to ignorance. When the test is set up this way, most of those tested get the right answers to a fairly high proportion of the items they reach. An arithmetic computation test of this type is included in most multiple aptitude batteries.

The ability measured by an aptitude test of this sort has been found important in many kinds of occupations that involve quantitative work--for instance, accountant, bank teller, business manager, machinist, bookkeeper, auditor.

Scores on this type of test do not have a particularly high correlation with scores on arithmetic reasoning, or other mathematics tests. A student may do well in mathematics in high school and college even if he is not

outstandingly fast and accurate in simple arithmetic computations.

Description of Talent Arithmetic Computation Test. Operations with both single-digit and multi-digit numbers are included. The four basic types of items are cycled (addition, subtraction, multiplication, division, addition.....). The division items all call for short division (no divisors greater than 12).

The operations tested are limited to relatively simple ones. Computation of square roots is not tested, nor is computation involving decimals, fractions, or percents. (These kinds of operations were omitted because ability on them has been found to be more closely related psychometrically to arithmetic reasoning and more advanced mathematics than to simple computation.) Sample items are shown below.

Item 123. Add:

423
33
29

A. 472
B. 475
C. 485
D. 575
E. 585

Item 124. Subtract:

403
-37

A. 340
B. 366
C. 376
D. 377
E. 476

Item 125. Multiply:

206
x94

A. 1824
B. 18024
C. 19364
D. 19464
E. 19564

Item 126. Divide:

3402 ÷ 6 =

A. 461
B. 531
C. 561
D. 567
E. 667

The answers are C, B, C, and D, respectively.)

The students are instructed to do their scratchwork in the test booklets, so that they do not have to copy the problems first. (This eliminates the effects that variable clerical speed and accuracy would have on the computation scores.)

Three scores are obtained for each student: number right (R), number attempted (A), and the composite (4R-3A). (Empirical evidence will eventually be obtained as to whether another formula might give better results for some purposes.)

The score given by the formula above (4R-3A) is designed to give an indication of the student's ability to perform simple computations accurately and at a high speed.

18. Table Reading

Purpose of test. The purpose of this test is to measure speed and accuracy of perception in a non-computational clerical task, involving obtaining information from tables. This kind of clerical aptitude is more complex than that measured by clerical checking tests.

Historical Antecedents. This type of test was developed during World War II, in the Air Force Aviation Psychology Program, as a "job sample" test for selecting navigators. The test was later found to be useful in predicting success in a wide variety of occupational activities, some of them not directly concerned with table reading at all. Table reading tests differ considerably in terms of the complexity of the task called for.

Among the kinds of jobs for which table reading tests are useful predictors are clerk-typist, draftsman, secretary, accountant, bookkeeper, statistician, statistical

clerk, shipping clerk, navigator. Of course the test is relevant to only one aspect of these jobs and therefore a high score on it does not guarantee success in the job, since other aptitudes and abilities are also involved.

Description of the Talent Table Reading Test. This test consists of a table with instructions as to how to use it, together with a number of multiple choice items that have to be answered by reading the table. The time allowance is very brief, to insure that nobody has time to finish.

Three scores are obtained for each student: number right (R), number attempted (A), and the composite (2R-A).

The task called for by this test, while not particularly hard, is sufficiently complex that slow learners may have a little difficulty figuring out what to do and how to go about it.

19. Clerical Checking

Purpose of Test. This test is designed to measure speed and accuracy of perception in a very simple clerical task involving working with verbal materials (names).

Historical Antecedents. Tests of this sort usually consist of pairs of names, numbers, groups of letters, or combinations; the subject is to indicate whether the two elements of the pair are exactly the same or different. These tests always have a very short time limit, so that no one will have time to finish. Ordinarily a very large number of items can be worked in a very short period of time.

These tests measure one of the very important aptitudes involved in clerical work. Tests of this sort have been

found relevant to success not only in a wide variety of clerical jobs but also in many other jobs which are essentially non-clerical in nature but which have a substantial component requiring this type of accuracy; for instance accountants.

A test of this type, on which the Talent Clerical Checking Test is modeled, is an important part of the United States Employment Service's General Aptitude Test Battery.

Description of Talent Clerical Checking Test. This test consists of pairs of names. The task is to compare the two names and mark S if they are exactly the same and D if they are different.

Below are three sample items:

Joe Abernathy	Joe Abernethy
Charles Q. Piltdown	Charles Q. Piltdown
Mary Ann Wilsson	Marianne Wilsson

(The answers are D, S, and D respectively.)

Three scores are obtained for each student: number right (R), number attempted (A), and the composite (4R-3A).

20. Object Inspection

Purpose of the Test. The purpose of this test is to measure speed and accuracy of perception of form. More specifically, it is intended to measure the ability to spot imperfections in small objects quickly and accurately, when comparing them visually with a standard.

Historical Antecedents. The type of ability measured by this test is useful in occupations where inspection is one of the job elements. It has been found to be important in many assembly-line jobs. It also has been found to play a part in some complex engineering activities where

inspection of detail is an important element.

Description of Talent Object Inspection Test. Each problem consists of pictures of five small objects, four of which are supposed to be identical. The remaining one is slightly different. The task is to pick out the picture that is different.

A sample item follows.

Item 130.

A

B

C

D

E

(The answer is D.)

Three scores are obtained for each student: number right (R), number attempted (A), and the composite (2R-A).

21. The Preferences Test

Purpose of test. This test has potentialities for being used in many ways. One posibility is that since it was given with an extremely stringent time limit, the mumber of problems on which the student has time to make his decision (i.e., to express a preference) will indicate his speed in making at least one sort of decision.

It is also expected that some empirical keys can be set up, which may be useful in other ways.

Historical Antecedents. The preferences test is one of those that was put in the battery for experimental purposes rather than because there are any clear antecedents which suggest that a test of this sort would be useful.

Description of Talent Preferences Test. Each item

consists of a pair of adjectives which might describe a person. The student is required to pick the adjective from each pair that he would rather have applicable to his friends.

The two characteristics paired are intended to be about equal in social acceptability, at least for an average group.

The Preferences Test is not intended to be a test of vocabulary. An effort was made to limit the adjectives to those that most high school students would understand. For this reason some highly colloquial terms were used (e.g., "smart-aleck", "stuck-up").

In considering the results derived from this test it is important to bear in mind that THE TEST IS EXPERIMENTAL. As yet nobody can be sure what it measures.

In regard to its relation to "decision-making", it must be remembered that only one, very limited, kind of decision-making is involved. While it is hoped that the measured trait will prove somewhat generalizable, it is recognized that this may not be the case, or that the trait may be generalizable to only a very limited extent. It is planned to investigate the hypothesis that at least under some circumstances, the test locates the individual on a continuum that has "snap-judgments" at one end and "indecision" at the other.

22. Themes

The two five-minute themes, described in Chapter 4, have been included in the battery for the following three reasons:

1. They will provide some insight into the student's values, attitudes, personality, motivations, and plans. (This is discussed in Chapter XI.)
2. They will provide a basis for determining whether the student can write coherent and correct English--free of mechanical errors (spelling, punctuation, grammar, etc.), well expressed, and logically organized. It is fully recognized, however, that themes as brief as these provide only limited scope for demonstrating the ability to produce a well organized composition.
3. In addition to their inherent value as indices of the ability to write effective English, they can be scored to provide an independent criterion of this ability, against which the relevant English subscores (primarily the Effective Expression score) can be validated*.

---

*To validate a test means to determine the degree to which it measures what it is intended to measure or predicts what it is intended to predict. The usual method of validation is to correlate the test with an appropriate independent criterion.

Table VII-1 Distribution of Information Test Items

	Number of Items		
	General Vocab.	Other	Total
Part I*			
1. Screening	-	12	12
2. Vocabulary	21	-	21
3. Literature	-	24	24
4. Music	-	13	13
5. Social Studies	-	24	24
6. Mathematics	-	23	23
7. Physical Science	-	18	18
8. Biological Science	-	11	11
9. Scientific Attitude	-	10	10
10. Aeronautics and Space	-	10	10
11. Electricity and Electronics	-	20	20
12. Mechanics	-	19	19
A. Tools, construction	-	(10)	(10)
B. Motors and mechanisms	-	(9)	(9)
13. Farming	-	12	12
A. Farm	-	(8)	(8)
B. Ranch	-	(4)	(4)
14. Home Economics	-	21	21
A. Cooking	-	(11)	(11)
B. Other	-	(10)	(10)
15. Sports	-	14	14
Subtotal	21	231	252
Part II*			
16. Art	1	11	12
17. Law	1	8	9
18. Medicine	1	8	9
19. Engineering	-	6	6
20. Architecture	-	6	6
21. Journalism	1	2	3
22. Foreign travel	-	5	5
23. Military	-	7	7
24. Accounting, business, sales	2	8	10
25. Practical knowledge	-	4	4
26. Clerical	-	3	3
27. Bible	-	15	15
28. Colors	-	3	3
29. Etiquette	-	2	2
30. Hunting	-	5	5
31. Fishing	-	5	5
32. Outdoor activities (other)	1	8	9
33. Photography	-	3	3
34. Games (sedentary)	-	5	5
35. Theater	-	6	6
36. Ballet	1	1	2
37. Foods	-	4	4
38. Miscellaneous	1	9	10
Subtotal	9	134	143
TOTAL	30	365	395

---

* Fifteen subscores and a total score were obtained on Part I. On Part II individual responses to each item were recorded on IBM cards.

References
Chapter VII
Background and Description of the Tests

(1) Bennett, G. K. Tests of mechanical comprehension. New York: Psychological Corp., 1940.

(2) Carroll, J. B. The factor analysis of two foreign language aptitude batteries. J. genet. Psychol., 1958, 59, 3-19.

(3) Carroll, J. B., & Sapon, S. M. Psi-Lambda foreign language aptitude battery. New York: Psychological Corp., 1955.

(4) Dailey, J. T., Lecznar, W. B., & Brokaw, L. B. Development of the airman classification test battery. Air Training Command Research and Development Program. Research Bulletin 48-4, 1948.

(5) Dailey, J. T., & Mays, R. M. Use of spelling and handwriting measures as a check on the technician specialty aptitude index. Air Force Human Resources Research Center. Research Bulletin 53-4, 1953.

(6) Dailey, J. T., Zaccaria, M. A., Foley, J., & Gerry, R. Electronics information test. San Antonio, Texas: Personnel Research Associates, 1958.

(7) Flanagan, J. C. Flanagan aptitude classification tests. Chicago: Science Research Associates, 1957.

(8) Guilford, J. P. (Ed.). AAF aviation psychology program research reports, No. 5: printed classification tests. Washington: U. S. Government Printing Office, 1947.

(9) Guilford, J. P. The structure of intellect. Psych. Bull., 1953, 267-293.

(10) Knoell, & Harris, C.

(11) Rice, J. M. The futility of the spelling grind. Forum, 1897, 23, 163, 409.

(12) Turse, P. L. Clerical aptitudes tests. New York: World Book Co., 1955.

CHAPTER VIII

THE BACKGROUND AND DESCRIPTION OF THE STUDENT ACTIVITIES INVENTORY, THE INTEREST INVENTORY, AND THE STUDENT INFORMATION BLANK

Part 1. The Student Activities Inventory

Early Considerations of the Test.

The Testing Panel at its meetings in May 1958, strongly recommended that a test of the "personality type" be included in the Talent battery of tests. It has been found repeatedly in many studies of talent and its utilization that personality factors are frequently an important reason why some persons fail to live up to their early promise. Panel members urged that self-report materials (those where the student answers questions about himself) be collected, supplemented by other material such as ratings by fellow students and by the teaching staff. The tremendous administrative difficulties involved in collecting such supplementary information were recognized, however, and the Panel's final recommendation covered only the use of a self-report inventory. As to the exact procedures to be followed, the Panel did not offer specific guidance.

Historical Background of Personality Testing

Personality assessment probably dates back to the earliest history of man, when one caveman estimated that he could bully another out of the last piece of mammoth meat. From that time to this, men have been trying to describe and predict the behavior of other men, for one purpose or another. The next few pages present an overview of the development of personality assessment, with emphasis on the type of personality assessment employed in Project Talent.

Definitions of Personality

Personality has been variously defined; Allport (1937, Ch. 2) has presented a very interesting discussion, in which the word is carried back to the Latin personna and traced through to present day English, giving a full 50 definitions in the process. However, a definition advanced by Woodworth (1929, p.553) appears appropriate to the survey nature of Project Talent and served to guide the development of ~~the~~ personality test:

> "Personality refers not to any particular sort of activity such as talking, remembering, thinking or loving, but an individual can reveal his personality in the way he does any of these things."

Early Beginnings in Personality Assessment

Alfred Binet. One of the earlier efforts to assess personality as such was made around the turn of the last century by Alfred Binet, whose inventive thinking resulted in the now famous Stanford-Binet Intelligence tests.

Binet (1903) describes one attempt to categorize psychological types by means of tests. The subjects of the experiment were his two daughters. The most important aspect

of the experiment consisted of asking his daughters to write lists of twenty words each. He then asked them to explain the meaning of each word and also the images that occured to them in connection with each word. Binet felt that this last step was important since sometimes an abstract word is used in a particular way (e.g. "home" meaning "<u>my</u> home"), and sometimes a specific word is used when an abstraction is meant.

When Binet analyzed the responses of his daughters he found Armande to be introverted and detached from reality, her images were vague and less precise than those of Marguerite. Marguerite's images were concise and vivid; Binet found her extroverted and practical.

In reaching these conclusions about his two daughters, Binet sorted the words they wrote into six categories: unexplained words, words naming objects which were present at the time, words referring to the person herself, words drawn from memories, words representing abstractions, and words taken from imagination. He noted that Armande's unexplained words (those for which she could give no reason as to why she gave the word) were more numerous than those of Marguerite, yet much less common. For example: messenger, character, professions, person; compared to: table, horse, plate, spoon, etc. Marguerite was classified by Binet as an "observer" type since she continually seemed to draw ideas for new words to write by looking at things around her. Marguerite named many things which belonged to her, Armande listed none. Binet felt that Marguerite was in fact much more concerned with and attached to her possessions than Armande.

In the case of abstractions, the two girls differed

considerably. Binet gives the example of listing the articles contained in a house saying: Marguerite visits a particular house mentally and names the objects by memory; Armande, on the contrary, thinks of any house and names the things that she knows exist in any house. Further, Marguerite produced no images which Binet could classify as imaginative (which he defined as detailed but non-existent), while Armande gave many. For example, in response to the words "road" and "sun" she described an imaginary road running beneath a burning sun, through an open field with no shadows.

Based on these experiments, Binet sought--but failed to find--a "ruling factor" or one-word description which would cover the total personality picture of each subject, as revealed by analysis of her responses.

The results of these experiments left Binet dissatisfied, particularly with his failure to be able to sum up the total personality picture in terms of one dominant aspect of it. Thus, some of the first attempts to assess personality scientifically only succeeded in emphasizing its diversity.

Carl Jung. Carl Jung devised his word association technique about 1906. This technique was oriented more toward personality assessment in abnormal people, though it supposedly applied to normals as well. In this technique, the subject was instructed by the examiner to respond quickly and without reservation, to each of a series of 100 words, giving the first word entering his mind in response to the stimulus word. A modified English translation of this list may be found in Jung (1919, p. vii). In the standard series of 100 words, a number of potentially emotional words (e.g. "love") were scattered among banal words (e.g. "box"). The examiner

noted each of the subject's responses, the amount of time elapsed between stimulus and response, and such behaviors as hesitation, embarrassment, and the giving of the antonyms. Inferences were then drawn concerning his personality.

As might be expected most of the associations given to such words were relatively innocent of psychological meaning. For example, responses such as "light--dark", "door--open", and "tooth--ache" were of little importance by themselves in judging personality. Similarily, many responses were of the synonym or rhyme types as "splendid--magnificent", "heart--part", etc. On the other hand, judgments of value seemed to offer important areas for interpretation, *e.g.*, "father--good", "ride--dangerous", and "small--unpleasant". Jung (1919) devoted considerable time to the statement and explanation of a detailed classification system of types of associative responses which is much too involved to describe here.

In his interpretations, Jung tabulated the number of each of the various kinds of responses given by his subject, using the classification system. Although he believed that even the most objective and trivial of associations had inner meaning, these were often much too obscure to evaluate. Thus, he concentrated on those associations having personal meanings. Some of these were unimportant, but taken together provided some insight into areas with which the subject is concerned. For example, the responses "schoolboy--soldier", "faithful--soldier", and "rank--file" form a group which tends to suggest the subject's concern with military service.

While this form of personality assessment found some favor particularly with the clinical and psychoanalytic

groups it required the individual attention of a highly trained person. Furthermore, relatively little was done in the way of objective validation of this procedure, though it became quite widely used.

Approaches to Personality Assessment

Arising out of the early beginnings discussed above and from the interest generated by the growing success of mental testing, many approaches to the assessment of personality evolved. Some idea of their variety can be had from those discussed by Allport (1937, Ch. 14), which are listed in Table VIII-1.

Table VIII-1

Approaches to Personality Assessment (Allport)

- Studies of Cultural Settings
- Physical records
- Social records
- Personal records
- Expressive movement
- Ratings
- Standardized tests
- Statistical analyses
- Miniature life-situations
- Laboratory experiments
- Prediction
- Depth-analysis
- Ideal types
- Synthetic methods

One approach not specifically mentioned above is important. Rorschach (1921) presented his ideas about personality testing, together with the outline for the Inkblot Test, in 1921. His invention of the test resulted from "an experimental investigation in form perception". In this investigation he found that manics, schizophrenics, and other abnormal groups had different characteristic ways of perceiving. He used inkblots since they contained no particular designs, and thereby allowed a maximum opportunity for the person tested to display his own personality in his

responses. The Inkblot Test is thus an investigation of a person's characteristic responses to neutral, standardized stimuli. Rorschach's methods of personality assessment are now considered part of a larger group of methods (projective techniques) all characterized by the opportunity for the subject to "project" his personality difficulties into different forms, situations and individuals.

Limitation of Approach.

Most of the methods and approaches to the study and evaluation of personality mentioned above are not adaptable to a large-scale survey. For this reason, among others, Project Talent adopted an inventory, self-report approach. The remainder of this review is accordingly limited to such instruments.

The Advent of the Inventory

Woodworth's Personal Data Sheet. Personality assessment was in a highly experimental state at the beginning of World War I. World War I had a profound effect on personality assessment. For the first time, the Armed Services were inducting men in wholesale numbers. A strong need soon arose for diagnosing the ability of those men to adjust themselves satisfactorily to the strains and stresses of military life.

Obviously, it was not possible for the number of psychologists available to interview personally every man drafted. The need was for a quick, paper-and-pencil type of screening device which could be administered in wholesale lots to draftees by relatively untrained personnel. It was to fill this need that R. S. Woodworth in 1917 developed the first of the group personality tests, his now famous Personal

Data Sheet (sometimes referred to in the psychological literature as the Woodworth Psychoneurotic Inventory).

Development of the Personal Data Sheet. The Personal Data Sheet (Woodworth, 1939,Pp. 18-19) was an attempt to standardize a psychiatric interview. In preparing the Data Sheet, Woodworth gathered information from psychiatric literature and from conferences with psychiatrists. From this information he constructed questions answerable by Yes or No. These questions were intended to reveal such symptoms as moodiness, guilt, and functional physical symptoms. The questions were generally of two types: those referring to the past history of the subject and those referring to his present.

Woodworth submitted the Data Sheet to standardization groups of "normal" individuals and known neurotic patients. No items answered unfavorably by 25% or more of the normal group were retained. The score made by a subject on the Data Sheet was taken to be the total number of unfavorable responses (sometimes Yes, sometimes No). If this score was excessive in comparison with that to be expected from the normal standardization group, neurotic symptoms were assumed to be indicated. A psychiatric interview was then arranged for the person.

Woodworth agreed that the Personal Data Sheet was only a screening device, but it was considered successful for this purpose. This test was the first widely used example of the questionnaire-inventory approach to personality assessment, and set the pattern for many similar questionnaires following the war.

Outgrowthsof the Personal Data Sheet. The impetus of the

war plus the success of Woodworth's Personal Data Sheet kicked off a strong spurt of enthusiasm for personality assessment. Many of the tests which resulted were direct outgrowths of the Personal Data Sheet, both in terms of form and content. Some of these tests introduced one or more changes in Woodworth's procedures, and most confined themselves to the measurement of "neuroticism", "emotional stability", "introversion-extroversion", and the like.

Some of the changes tried out during this period were the use of a graphic scale (Laird, 1925), where the subject indicated his choice by placing a check mark on a line instead of answering "Yes", "No" or "?". Another (Heidbreder, 1926,'27) introduced the idea of selecting items by measuring how well they predicted ratings of the subjects made by themselves and their friends. In another, G. W. and F. H. Allport (1924, 1928) experimented with the use of items composed of descriptions instead of statements (for example: "At a reception or tea do you seek to meet the most important person present?"). In addition, they developed an important new idea in scoring--giving different items different weights in determining the total score.

Another important innovation was made by L. L. Thurstone in his 1928 Personality Schedule. This Schedule consisted of items designed to measure "Neurotic Tendency". Thurstone decided which answers were to be considered symptomatic of neurotic tendency. The Schedule was given to 694 freshman college students. Scores were obtained by counting the questions to which "neurotic" answers were given. Thurstone then selected the 50 most and 50 least neurotic-scoring individuals. He then went through the test item by item,

counting for each item the actual number of "neurotic" answers given by each group. He justified his original assignment of scoring weights to the items by noting that the high scoring group gave "neurotic" answers more frequently than the low scoring groups. No external standard of "neuroticism" was used. Thus, items were retained in the Thurstone inventory on the basis that they agreed well with his *a priori* keying. Such "internal consistency" methods have since been widely applied in the development of personality inventories.

An additional step in the development of one-dimensional personality assessment is represented by the Terman-Miles Attitude-Interest Analysis Masculinity-Femininty Test (1936). An external criterion was used for the determination of item significance as in the case of Heidbreder and the Allports; however, in this case, the criterion groups were selected in advance on a completely external basis, and the test scores had no part in their selection. Thus, the measurement approach was to select items to differentiate these predetermined groups rather than to select items to measure a visualized trait and then to define criterion groups in terms of the visualized trait.

Summary. At this point, it appears profitable to pause and briefly summarize the developments in personality assessment from the historical Woodworth Personal Data Sheet to about 1930. During this period, the success of the Personal Data Sheet as a quick, relatively satisfactory device for screening military personnel for neurotic instability led to a boom in personality assessment. Literally hundreds of tests were developed; only a few of these measures have been

mentioned.

This period saw the rise and fall of measures of personality of many kinds, including the first of the major projective instruments. The period was characterized by a large number of self-report inventories, more or less traceable to the Personal Data Sheet. Virtually all of the inventories of this period were single dimensional in character, measuring such traits as "Neurotic Tendency", Introversion-Extroversion", and "Ascendance-Submission". Numerous variations of the "Yes-No" type of response were tried out, including graphic rating scales, algebraic summation of unit item weights and algebraic summation of differential item weights.

THE ORIGINAL DOCUMENT WAS OF POOR QUALITY. BEST POSSIBLE REPRODUCTION FROM COPY FURNISHED ASTIA

Particularly important was the emergence of a number of distinct procedures for item selection. These included the following major types:

a. Judgment.
For example, items were included in Woodworth's Personal Data Sheet primarily on the basis of his judgment and that of psychiatrists.

b. External Correlation.
For example, Heidbreder and the Allports correlated their items with external criteria, namely the ratings of the subject's associates.

c. Internal Consistency.
For example, Thurstone related his items to total score on his Neurotic Inventory by comparing the number of neurotic answers for the item for those scoring high on the test to the number of neurotic answers for the items for those scoring low on the test.

<u>Multidimensional vs. Unidimensional Inventories.</u> Unidimensional tests dealt with measurement along one defined dimension or trait at a time, *e.g.*, introversion-extroversion. While many such tests are in use today, they are often reserved for special purposes, *e.g.*, Wallen's (1943) Food

Aversion Test, Taylor's (1953) Anxiety Scale, Cassel's (1952) Level of Aspiration Test, and Maslow's (1945) Security-Insecurity Test. Most personality inventories today are of the multidimensional type.

In some cases, a multidimensional inventory is simply a group of unidimensional inventories under the same cover. In other cases it may contain an entirely different approach to test construction, such as the one Bernreuter (see below) used.

The unidimensional tests discussed in the previous section were widely used. However some investigators soon began to search for more efficient ways of assessing personality since the unidimensional approach either focused attention on one aspect of the subject's personality, or else required the administration of a separate test for each trait in question.

The unidimensional tests were based upon the assumption that a given behavior was related to one trait. The new type of multidimensional inventory was an attempt to test the assumption that a given behavioral element (an item) could be related to several different traits, and thus might appear, differently weighted, in more than one scale of measurement. This assumption, if supported, could make possible the construction of tests which would simultaneously measure several traits with the same items.

The Bernreuter Personality Inventory

The Original Inventory.

The Bernreuter Personality Inventory was the first one of this latter kind of multidimensional inventory. The development has been described in detail by various authors

including Ferguson (1952). A discussion of the theory underlying this inventory is available in Bernreuter (1933).

Bernreuter's Personality Inventory was first published in 1932. It consists of 125 items answerable by "Yes" or "No", drawn from four earlier unidimensional tests: Laird's C-2 Introversion Test, the Allport's Ascendance-Submission Reaction Test, Thurstone's Neurotic Inventory and Bernreuter's own Self-Sufficiency Test. The Inventory was intended to measure simultaneously the four dimensions measured by these tests, and contained four scales called Neurotic Tendency, Self-Sufficiency, Introversion-Extroversion, and Dominance-Submission.

After the items were assembled, Bernreuter gave a trial form of the Inventory together with the four original tests, to about four hundred students from various universities. Using these scores, Bernreuter selected groups to represent extreme scores on each of the tests. For each trait measured by the Inventory, there were two groups: 50 high-scorers (25 men, 25 women) and 50 low scorers (25 men, 25 women). Bernreuter then computed scoring weights for each question on the Inventory. Different weights were assigned to the same items for different traits, depending upon the extent to which the item differentiated the high and low score groups for that trait. The subject's score on a given scale is the algebraic sum of the weights for his answers in that scale.

Several items in the Inventory are given below, along with, for example, their scoring weights on two of the scales:

	Neu.Tend. (B1-N)			Self-Suff. (B2-S)		
	Yes	No	?	Yes	No	?
Are you troubled with shyness?	6	-7	1	-1	1	-1
Have books been more entertaining to you than companions?	3	-2	-1	2	-3	1
Do you prefer making hurried decisions alone?	-1	1	-2	6	-6	-1

Bernreuter found generally satisfactory reliability coefficients for his scales (split-half coefficients ranging from .82 to .91). In addition, he demonstrated that his four test scales correlated virtually perfectly (when corrected for attenuation) with the four tests they were supposed to replace, respectively. It therefore seems clear that Bernreuter's original idea that one set of items can be weighted differently and can serve several purposes at one time is well-supported. However, the demonstration of relationship with the original four tests, is still not a good demonstration of the validity of the Inventory, since the original four tests have only modest validity.

The Flanagan Revision.

Flanagan's (1935) factor analysis of Bernreuter's Personality Inventory was a pioneer effort in the application of a new statistical technique--factor analysis--to the area of personality testing. Struck by the high intercorrelation of Bernreuter's scales, Flanagan's objective was to replace the arbitrary construction of the trait definitions with a method for scientifically obtaining significant traits. Bernreuter had selected four personality traits, and provided scores on each. Flanagan gave the Inventory to 305 adolescent boys and obtained the intercorrelations shown below:

		B2-S	B3-I	B4-D
Neurotic Tendency	(B1-N)	-.39	.87	-.69
Self-Sufficiency	(B2-S)		-.3	.51
Introversion	(B3-I)			-.62
Dominance	(B4-D)			

He then applied Hotelling's principal components factor analysis, and found that two dimensions rather than four, were able to account for virtually all of the individual differences in test scores.

Flanagan then re-worked his data, and in trying his new keys out on groups of children, he found good reliabilities (.86 and .78) and a very low interrelationship (.04) between his new dimensions, showing that they represented two really different aspects of personality.

Flanagan's work resulted in the addition of two scales to the original Bernreuter, one of which he called Self-Confidence (F1-C) and one which he called Sociability (F2-S). The present edition of the test contains all six scales, although Flanagan's two scales may be used alone instead of the original four scales.

Flanagan's work was historic in its application of factor analysis to the personality inventory. Nevertheless, he worked with the original Bernreuter material and of course his scales are subject to the limitations contained in that material.

Further Factor Analytic Developments

Factor-based Tests.

Flanagan's success at applying factor analysis to personality measurements initiated a widespread interest in this technique, and a number of the instruments developed in recent years have been based upon a factor analytic approach.

Those in favor of factor analytic approaches have argued that factor analysis permits the description of human behavior in terms of a relatively few, relatively independent, more basic dimensions. Several notable examples of the use of factor analysis in this area are mentioned.

The Guilford Inventories. These inventories, developed from 1936-1939, represent extensive use of factor analysis. Whereas Flanagan's application of factor analysis to the Bernreuter was based on the relationships (intercorrelations) among scores on the four scales, Guilford analyzed the relationships among items. This process enabled Guilford to define an estimated number of related factors, rather than having to start with a number of scales developed by theorizing and then refining these into a number of factors. Thus Guilford used factor analysis as a basic scale construction technique, assigning item scoring weights in rough agreement with their importance to the factors (factor loadings).

The Guilford Inventories cover 13 traits in three Inventories as follows:

a) An Inventory of Factors STDCR (Social introversion-extroversion, thinking introversion-extroversion, depression, cycloid disposition, and rhathymia or carefree disposition).
b) The Guilford-Martin Personal Inventory (objectivity, cooperativeness, and agreeableness).
c) An Inventory of Factor GAMIN (general activity, ascendance-submission, masculinity-femininity, inferiority feelings and nervousness).

An Inventory of Factors STDCR was developed basically from a number of sources, including Jung's writings, Freyd's monograph (1924), and several earlier introversion-extroversion inventories. Centroid factor analysis of the intercorrelations of the items, based on the scores of 610 men

and 390 women, formed the basis for the initial form of the inventory. The development of the final form is not exactly described.

An Inventory of Factors GAMIN was designed in much the same way to test the idea of an activity drive. It was also supplemented by material from other studies. The Guilford-Martin Personnel Inventory followed the same general pattern of development, and was intended to assist in the identification of maladjusted workers and to extend the list of traits already assessed by the other inventories.

Reliability coefficients for the 13 factors are good (from .80 - .94) and some, though not a great deal, of evidence has been collected as to the validity of the inventories.

The Guilford Inventories demonstrated among other things the utility of factor analysis in scale construction and in clarification of old concepts. For example, the development of the Inventory of Factors STDCR clearly suggested that the old idea of introversion-extroversion is not a single trait, but is composed of at least five sub-traits.

<u>The Thurstone Temperament Schedule</u>. The interrelationships among the original 13 factors isolated by Guilford were quite high. This fact led a number of researchers to re-analyze Guilford data. Thurstone (1951) carried out such a re-analysis and concluded that seven major factors were sufficient to account for virtually all of the information. These seven traits were described by Thurstone (1950) as follows:

Active (A) -- Working rapidly, even when not necessary; restless.

Vigorous (V) -- Physically very active using large muscles and much energy; outdoors, sports, manual work activities.

Impulsive (I) -- Quick decisions, carefree, happy-go lucky.

Dominant (D) -- Leadership initiative, responsible.

Stable (E) -- Cheerful, even disposition.

Sociable (S) -- Enjoy company of others, make friends easily, sympathetic, cooperative.

Reflective (R) -- Quiet, work alone, meditative, theoretical.

The reliability coefficients of the scales are generally regarded as relatively low (ranging from .45 - .86 for different scales in different studies). In addition, the interrelationships of Thurstone's seven primary factors, were surprisingly high (in two cases .71 and .52), considering the techniques used. Validity studies involving this schedule are relatively scarce, and many reviewers recommend its use for research purposes only.

Cattell Tests. Some of the most extensive quantitative research yet done in the area of personality evaluation has been done by Cattell (1946, 1957). In his effort to develop a comprehensive description of personality, Cattell started with the Allport and Odbert (1936) list of almost 18,000 trait names. He reduced this to 171 by combining closely related concepts, and began his factor analytic studies of these traits.

A number of important points about Cattell's work must be mentioned. First, he has set up a systematic plan for investigating the area of personality assessment; second, he has rejected the idea of traits taken by themselves and has insisted that they must be dealt with in terms of the conditions or situation immediately present; and finally, he has distinguished between "surface" traits and "source" traits, the latter corresponding to factors which give rise

to the surface traits but can only be discovered by mathematical analysis and psychological insight.

As might be expected, Cattell has devised a number of tests to measure the factors growing out of his studies. Two examples are:

a) The Sixteen Personality Factor Questionnaire. The inventory was constructed to measure 12 of the source traits found in Cattell's early factorial studies and four additional ones obtained from the study of questionnaire responses. Questions are of the "Yes", "In between", or "No" type. Good reliability coefficients (split-half, in the .71 to .93 range, based on 450 cases) were reported. Considerable validational work is underway. The manual is not clear on the developmental details of the test, but reviewers have agreed that a large amount of statistical work has gone into the test, and that it is likely to prove a significant step in the assessment of personality.

b) The Objective-Analytic Personality Test Batteries. Published by Cattell, et. al., in 1955, these tests consist of one adult battery designed to cover 18 of Cattell's source trait factors (Cattell's Universal Index Numbers 16-33) by 43 group tests or 54 individual tests (also available in abbreviated forms for 12 of these factors, and for single factor batteries).

These tests require that the user be thoroughly grounded in Cattell's complex system of concepts, terms, and theories. In addition to the complexity of the theoretical system involved, however, the average test user will find little evidence as to the relationship of these batteries to clinical, industrial, and other outside criteria, no norms, and no

data on reliability. Thus, while the promise of these batteries may be great, they are primarily restricted to the researcher at the present time.

Summary. The work listed above, as well as that of other researchers, illustrates the variety of steps being taken through the techniques of factor analysis to attempt to better define and measure personality. In spite of the experimental, highly complex, and mechanistic nature of the methodology employed, considerable progress appears to have been made with more expected in the near future.

Non-Factor Multidimensional Inventories

General

Although the Flanagan treatment of Bernreuter's Personality Inventory was a significant step, the development of multidimensional inventories was not confined to those based on factor analysis. Some of the more prominent of the early multidimensional, non-factor tests are mentioned below.

The Bell Adjustment Inventory. This widely used inventory has two forms, one for high school and college students (prepared in 1934) and another for adults (1938). The adjustment areas measured by the inventory are: home, health, emotional and occupational. The inventory is self-administering, and the score for each area is the number of significant answers marked by the subject. Area scores can be summed to obtain a total adjustment score for men and women in high school, college, and at the adult levels.

The test distinguished well between two groups of adults classified as very well and very poorly adjusted by counseling experts. Good reliability coefficients have been

reported (from .81 to .94). Its chief use is as a large-scale screening device since the assumption that each question contributes to only one adjustment area is questionable, and the various areas appear to be highly related.

C. <u>The Minnesota Multiphasic Personality Inventory (MMPI)</u>. One of the most widely used of personality tests, the MMPI was developed in 1940 by S. R. Hathaway and J. C. McKinley for the purpose of psychiatric evaluation. The inventory consists of more than 500 statements to be answered "True", "False" and "Cannot say". Items were selected for the final test only if they showed a large percentage frequency difference between a criterion patient group and a normal group. The clinical scales in general use are: hypochondriasis, depression, hysteria, psychasthenia, schizophrenia, hypomania, and social introversion-extroversion. Test-retest reliability coefficients for these scales range from .46 to .93. There are separate norms for men and women, although a number of investigations have indicated that these norms are too general, and need work.

One of the most important features of the MMPI is the attempt to provide control scales to indicate the attitude of the subject taking the test. The "?" scale indicates the extent of omissions in responses; the "L" (lie) score measures the degree to which social desirability is influencing the responses; the "F" (validity) score is designed to identify indiscriminate markings, carelessness, lack of understanding, etc.; and the "K" (correction) scale is believed to indicate the person's test-taking attitude--defensive, rigid, self-critical, etc. While these four scales have not been completely successful, they have contributed

greatly to the success of the MMPI.

For its purpose (psychiatric evaluation), the MMPI is generally considered one of the most valid of available questionnaire instruments.

Forced Choice

The general self-report inventory, even though it might be scored for several dimensions, came in for considerable criticism in several respects. Many studies have demonstrated that people could present very unreal views of themselves when this seemed desirable to them. This situation is particularly bad where a person may have something to gain by presenting a false picture of himself, as in industry or the military. In other cases people have experienced great difficulty in choosing "Yes", "No", or "?" as an answer to a question. In response to these criticisms, several psychologists, notably Jurgensen (1944), Sisson (1948), and Shipley (1946), developed the "forced choice technique". Essentially, the subject is required to choose between two alternative statements or answers which appear equally desirable or acceptable, but which differ with respect to their scoring weights for the trait in question.

Where the paired alternatives are effectively equated for social acceptability and apparent desirability for the group in question, the opportunity for faking is reduced. In addition, by requiring a relative rather than an absolute judgment, the forced choice item tends to reduce the ambiguity usually associated with items such as "often", "rarely", etc.

In using the forced choice technique, fakability and ambiguity are reduced but not eliminated. Whenever the person

who takes the test has a strong reason to present himself in a particular light, he can usually do so to some extent, depending on how obvious the relationships are of the statements in the item to the traits being measured. Where there is little reason to fake, the traditional self-report type of inventory may be preferable by virtue of its greater simplicity and understandability.

Recent Inventories

This section summarizes a few recent multi-dimensional inventories. A considerable number of single-dimension inventories are still being produced for special purposes, but in most cases have not found particularly widespread usage. While factor-based batteries appear from time to time and forced choice form is quite popular, the traditional trait-oriented, self-report inventories are still used in large numbers. It should be noted that the Project Talent test made no effort to include all the areas covered by the tests described below. The tests reviewed here simply illustrate some of the directions presently being taken by test producers in their attempts to assess personality by means of self-report inventories.

California Test of Personality. (Thorpe, Clark and Tiegs). This inventory, originally published in 1942, was revised considerably in 1953, at least partly as a result of fairly severe criticisms by reviewers. The test is of the traditional "Yes--No" type, and is really a battery of five tests for different age levels. It assesses two general areas: Personal Adjustment includes self-reliance, sense of personal worth, sense of personal freedom, feeling of belonging and freedom from withdrawing tendencies and

nervous symptoms; Social Adjustment includes social standards, social skills, freedom from antisocial tendencies, family relations, school (or occupational) relations, and community relations. The questions in the test attempt to elicit information indirectly. Rather than demanding "Do you hate people?" the test asks, "Are certain people so unreasonable that you hate them?" The authors feel this enables the subject to "save face" and thus encourages him to answer more freely. This test assumes that the subject is constantly adjusting to problems and attempts to evaluate his adjustment.

While the test has undergone much internal consistency and other statistical work designed to strengthen its validity and reliability, it still possesses some defects. The reliability coefficients of the two part scores and the total score (alternate from computation) are .90, .89, and .93, respectively. However, diagnostic or profile interpretations based on the sub-test scores are questionable due to the somewhat lower reliability of these sub-test scores. The intercorrelations of the various scores are also quite high. New validity studies based on the 1953 revisions are reported in the manual and the results are moderately favorable.

The California Psychological Inventory (Gough). One of the most recent inventories, published in 1957, this test was designed to deal with widely applicable characteristics of human behavior, and emphasizes positive and favorable aspects of personality rather than abnormality. The test is composed of 480 "True--False" items, many derived from the MMPI, ranging over an exceptionally wide content.

Eleven scales were derived from external criteria, and four from items judged to be homogeneous. Finally, three "control" keys were developed, similar in purpose to those developed for the MMPI: Sense of Well-being (Wb), based on responses given by normals told to present a bad picture of themselves; Good impression (Gi), based on responses of normals told to present good impression; and Communality (Ca), the highly popular responses. A marked deviation on a control score casts doubt on the individual's other scores.

While the manual is not complete with respect to all details of scale development, the procedure was essentially to:a) assemble a large stock of significant-appearing personality items; b) identify criterion groups differing sharply on some dimensions considered socially and/or psychologically significant; and c) develop a scoring key composed of those items found empirically to differentiate the criterion groups. This procedure may reduce "fakability", and may result in efficient keys, but also may result in a series of dimensions whose meaning is not clear.

The test-retest reliability coefficients are relatively low, ranging from .38 to .75 for one study, and .49 to .87 for another. Considerable validity information is available for such a new test, but much of it is seriously misleading, being based largely on inappropriate statistical procedures.

The Edwards Personal Preference Schedule. This test, published in 1953, manual revised in 1957, is notable in several respects. First, it is based solidly upon psychological theory. Edwards selected 15 of the Murray (1938) psychological needs as the dimensions to be measured in his test.

Second, in trying out an early version of the test in "True-False" form, Edwards found that the frequency of endorsement of his items correlated very well (.87) with their "social desirability" (as scaled by a group of 152 judges). In order to avoid having test scores determined primarily by social desirability, Edwards adopted a "forced choice" form in which items consisted of pairs of statements from the different scales, equated for scale values in social desirability. The final form consisted of 225 items wherein statements from each of the 15 scales were paired off twice against statements from each of the other 14, and 15 pairs were repeated in order to get a consistency score.

Acceptable reliability coefficients are reported (.60 to .87 split-half, corrected; and .74 to .88 test-retest at a one week interval). A number of validity studies are reported in the manual, some of which include a modest validity for some of the test scores. However, a study by Corah, Feldman, _et al_., found highly significant differences in social desirability between paired statements presumably matched for it, and thus concluded that Edwards had _not_ been very successful in controlling social desirability. Even so, the test is widely considered to be a very promising instrument.

The SRA Youth Inventory (Remmers, Shimberg, and Drucker). Published in 1949, manual revised and technical supplement added in 1953, this inventory is somewhat different from the ones discussed above in that it is geared specifically to the assessment of "problems" of high school youths. The 298 items were derived from a content analysis of "hundreds of essays" written anonymously by high school

students about their problems. This analysis resulted in a division of the problems of high school students into the eight areas covered by the inventory: 1) My school, 2) Looking ahead (to college, work, etc.), 3) About myself (personal adjustment), 4) Getting along with others (social adjustment), 5) My home and family, 6) Boy meets girl (dating, marriage, etc.), 7) Health, 8) Things in general (world affairs, religion, social ethics, etc.). Students are asked to check all statements which concern one of their problems. Scores are simply the number of items checked in each area.

A 9th score may be obtained, called a Basic Difficulty score. This consists of the number marked of the 101 items that all or 6 of 7 experts in the fields of psychology, guidance, and education agreed were indicative of basic personality difficulties.

Problem statements were classified into the 8 areas judgmentally, and these assignments were checked by internal consistency procedures. The technical supplement reports the 8 areas as "fairly homogeneous" with respect to both statistics and face validity, but the statistical criterion is not presented.

Reliability coefficients (only 48 cases) were good as computed by test-retest methods (.72 to .88) for the eight areas and the Basic Difficulty score. Several external validation studies are reported in the Technical Supplement which indicate some modest validity for the Basic Difficulty score in identifying adjustment problems, especially for girls. The discrepancies in results of self-rating and rating by others are clearly brought out, however.

The normative sample is national and well-described.

The authors suggest the use of the inventory as an adjunct to the counseling interview and it appears that for this purpose it should be quite useful in the hands of professional psychologists and counselors.

d. Critique and Evaluation

From almost the beginning, paper and pencil personality tests have been under criticism of one form or another. Over the years persons such as Allport (1924), Garrett (1933), and Ellis (1946, 1953) have voiced general criticisms of a consistent nature. Ellis (1946) has summarized the situation, including common objections and strengths of personality questionnaires, based upon his review of 259 studies.

Of the 259 studies, 80 showed positive validity, 44 questionable validity and 135 no validity for the questionnaire being studied. Internal consistency measures were not considered measures of validity. Of the questionnaires considered in these studies, only the MMPI showed up well with 8 positive validity studies out of 13 reviewed. The Bernreuter Personality Inventory was a poor second. It should be noted however that Ellis' criteria were unusually severe (he defined correlations of 0 - .39 as mainly negative, .40-.69 as questionably positive and .70 up as mainly positive). However, in his 1953 review of studies done since 1946, Ellis found that more research and clinical use was then being made of personality inventories than of any other type of assessment instrument.

The problems of paper and pencil personality testing may be summed up in brief as follows: Reliability data are in most cases adequate, but problems of validation run through most of the tests. The main issues in the self-report type

of inventory center around three areas:

1.) <u>Frankness.</u> Most subjects can successfully follow instructions to fake their responses in some particular way, even those who can't, often present a distorted picture of themselves.

2.) <u>Self-insight.</u> In order to be able to respond with valid information about himself, the subject must of course be aware of his true nature. Studies have repeatedly shown, and personal experience often bears this out, that people often aren't aware of their personal quirks and defects, and thus can hardly be expected to report them in response to a question.

3.) <u>Reading Load.</u> Inventories often run to several hundred questions in their efforts to obtain acceptable validity and reliability. A poor reader, after struggling with this for a while, may give up in his attempt to understand each question and answer an unpredictable number of them superficially, if not completely randomly.

While the reputation of personality questionnaires apparently reached a low point with the publication of Ellis' 1946 review, and while Super (1949, p.527) has said that perhaps the major trend since 1930 has been away from the inventory technique and toward various projective devices, the personality inventory is far from dead. Indeed, with the work of the factor analysts, and the development of a number of refinements such as forced choice form, control scales, and attempts to nullify the effects of social desirability, there has been a great deal of renewed interest in the inventory technique.

On the positive side of the ledger, the personality inventory may be thought of as a formalized interview with the great practical advantages of objectivity and standardization, low cost per individual case, ease of scoring and administration, and relatively low demands on the time of both subjects and highly trained personnel. In spite of all of the problems in obtaining satisfactory predictive validity, many studies show a moderate level of validity for inventories in differentiating between various groups. Also, there is plentiful evidence which indicates the existence of consistencies in human behavior which may be susceptible to measurement of the inventory type. Finally, it is clear that individual attention is a luxury, and that large scale military, industrial, and school programs will continue to demand some personality assessment at least at the screening level, using paper and pencil inventory techniques. Thus, it appears that a place for such techniques is assured, and that the need is to refine and improve such techniques to provide more valid results. Efforts at such improvements are currently being made in many directions. Refinements in technique and theory such as those discussed earlier, coupled with a re-examination of validation procedures and continuing search for new and more appropriate validity criteria, appear to hold considerable promise of progress in this area.

Resolution of the Problem

In measuring personality in the present study, the recommendations of the testing Panel and the conclusions suggested by historical and developmental considerations were adapted to the specific requirements of the problem at hand.

Thus, the present study employed personality measures

which were essentially self-administering, not very time consuming, and adaptable to machine scoring and computer analysis. In general then the test was to be one which fairly represented the "state-of-the-art" in assessing the personality of normal individuals, yet was tailored to the needs and purposes of the project.

Purpose of the Test

In addition to possible contributions to measurement techniques and knowledge about personality in general, the personality inventory was intended to provide information bearing on the several project goals. Examples of possible uses of the personality inventory information included the study of: 1. any relationship between high school policies, conditions, etc., and personality variables; _e.g._, if enrichment produces better college grades, is this equally so for those with equal ability but different personality patterns; 2. similarly, the effect of college policies and practices; 3. relationship of personality variables to the other factors which determine the effective utilization of the national potential; 4. college and career plans by personality trait, _e.g._, do highly dominant people plan careers as executives more often than less dominant people; 5. personality traits as related to decision-making; to levels of aspirations, etc.

It should be noted that the Project Talent test made _no_ attempt to measure abnormalties or psychiatric problems. On the contrary, it was intended to be used only with normal individuals in an effort to provide additional information about the way that people use their talents and abilities.

In designing the test, one point to be considered was the type of the test to be used. Of course, it had to be

multi-dimensional, but should it be a collection of uni-dimensional tests (where each item is scored for just one scale), or should it be of the Bernreuter type? The latter type is more efficient since several scores may be gotten from the same set of items using different weights for the items in different scores. This efficiency is somewhat offset however, but the fact that the scores obtained from such a test may show some interrelationships which are artificial; that is, they come from the fact that the <u>same items</u> are scored for <u>different traits</u>, not necessarily from the interrelationships of the traits themselves.

Flanagan (1935a,1935b) has shown that such artificial relationships are often of little importance, and that statistical corrections can be applied to eliminate themselves. Nevertheless they introduce complications, and it was decided to score each item for just one trait.

Measurement Approach

It is commonly agreed among most investigators that there are at least several reasons for the limited successes attained in the field of personality measurement. Among these are: 1) the inaccessibility (within the person, of the basic determiners of behaviors, thus requiring indirect measurement and inference; 2) the apparent complexity of the determination of any bit of behavior, involving the interaction of many needs, vast funds of personal experience, and constantly changing external conditions; and 3) the fact that various conditions often result in similar behavior, and, conversely, that similar conditions often give rise to vastly different observable behaviors.

Under such conditions, the trait approach, while descriptive of behavior, may often yield a surface type of measurement, lumping together (or separating) behaviors arising from vastly different (or highly similar) internal, motivational conditions. On the other hand, the factor analytic approach, while it may introduce a condensation of data, may also lump together a great variety of constructs, dynamically different in origin. In addition, it is often difficult to translate factor analytic data into reliable and real behavioral terms.

In view of the limitations of both of these approaches it was decided to combine some of the features of both the factor analytic and the trait approaches. In brief, previous work plus careful rational analysis formed the foundation for defining a number of narrow, but relatively pure, aspects of high school behavior. These were not designed to cover all high school behavior, but to sample selectively at points well distributed throughout the range of high school behavior. This procedure was designed to result in relatively good measurement of such narrow traits. Assuming that the underlying broad factors of personality consist of highly interrelated aspects of personality, and that the defined narrow traits represent some of many such aspects, it then appears reasonable that measurements or selected combinations of measurements of such narrow traits should be quite well related to the broad factor concerned.

Development of the Test

In taking the narrow trait approach it is important that these narrow traits sample much of the important aspects of normal high school behavior of a personality nature. Several

steps were taken to insure such coverage:

A. Criteria. Narrow traits were derived from a consideration of high school behaviors according to several criteria.

1. The behavior occurs in high school and is observable.
2. It can be well-enough defined to be identified by a high school student.
3. It is reasonable to expect a high school student to be able to rate himself on this behavior.
4. The behavior can be expected to be importantly related to future behavior.

There are many important aspects of normal behavior which were eliminated according to one or more of these criteria as being less than maximally important for this study.

B. Procedure. Based upon literature review, psychological analysis, and the application of the above criteria, a set of narrow traits was defined. It was felt that the essence of the narrow trait approach lay in the selection of several behavioral adjectives which would define the behavior indicated by the trait. Several approaches to the selection of such behavioral descriptions were employed.

1. Rational Analysis. Each of three investigators independently set down a series of narrow traits, defined by descriptive behavioral adjectives, derived from thinking about what they and their own high school acquaintances had been like. General agreement among these three independent sets of narrow traits was surprisingly good.

2.) Adjectival Analysis. The entire domain of effective behavioral adjectives was examined using Roget's Thesaurus as a source. The above criteria were applied to 2000 adjectives and the remaining 200 words categorized into a system of narrow traits. To insure complete coverage a 5% sample of Allport and Odbert's (1936) exhaustive list of 17,953 trait names was examined. Correspondence between the Roget and Allport coverage was good. The adjectival categories also agreed well with those devised earlier.

3. Relationship between Narrow Traits and Broad, Higher-order Factors. In order to provide some tie-in factor studies, reviews of the literature were studied, including French's (1953) critique. Such reviews provided information through which the narrow traits were logically related to the major second-order factors found most consistently in more than 70 studies. These six factors were: Dominance, Sociability, Drive, Self-Sufficiency, Masculinity, and Mature Personality.

The narrow traits derived through the processes described above cover many of the aspects of behavior formerly classified under one or another of the six higher-order factors, and at least some of the aspects of behavior from each of them. Moreover, the behaviors falling under them are more clear-cut, more easily definable and easier to measure by self-rating on the high school level. Some of the traits are more closely related to some of the higher-order factors than others, but most are related to one or another of the six.

C. Narrow Traits Tried Out. Consideration of all of the available materials, lists and background led to certain

recommendations for traits to be measured.

The scales listed in Table VIII-2 below included all of those prepared for the tryout of the preliminary form of the test. A number of these were omitted from the final form on the basis of tryout results, limitations of testing time, or booklet space. A description of each scale may be found in Appendix 8a.

Table VIII-2

<u>SAI Preliminary Tryout Scales</u>

Leadership
Vigor
Productivity
Tidiness
Sociability
Self-Confidence
Cheerfulness
Culture
Talkativeness
Impulsiveness
Persistence
Calmness
Social Adjustment
Theoreticality
Responsibility
Social Sensitivity
Group-Centeredness
Conventionalism

<u>Format of the Experimental Test</u>. The preliminary form of the personality test had about 15 to 20 items for each of the 18 scales tried out. It was anticipated that tryout data would make it possible to come out with about 10 scales of about 15 items each for the final form.

The test was divided into several parts. Parts I and II were short statement items concerning "What I do and the way I do it." Part I contained items best answered by options describing how <u>often</u> a statement was true, and Part II contained items best answered by how <u>well</u> the statement

described the student. Of course, many items fitted in either part and these were simply placed in either part.

Sample Format:

Regarding the things I do and the way I do them,
this statement describes me

A. extremely well. (almost always,)
B. quite well. (often.)
C. fairly well. (about half the time,)
D. not so well. (not very often.)
E. poorly or not at all. (almost never.)

Item 1. I have planned my own future.

Item 2. I like to be by myself.

Part III consisted of statements about "How other people describe me," *i.e.*, "Regarding the things I do and the way I do them, this statement describes what other people say about me." etc. Part IV consisted of a number of words, rather than statements, "that describe me". Part V also consisted of "Words that describe me", given in the same form as Part III (words other people would use). There were a total of 300 items in the experimental form.

The Administration of the Experimental Test, Fall 1959. The experimental form of the test was given to several groups, each of which had taken a well-known personality test previously. In this way, it was possible to get item analysis data and to estimate the relationships of the experimental scales with other tests at the same time. The details of this tryout are presented more fully in Appendix 8a.

The Final Form. On the basis of the tryout data, several scales were omitted, numerous items dropped, and some scales combined. This resulted in a final form of 150 items and 13 scales.

It is not expected that the 13 scales scored in the final form will cover the full range of high school personality. They represent, rather, the best-measured of the initial group of scales. However, efforts will be made to retain the much broader coverage of the 150 individual items themselves, if feasible, to relate to student behavior. Thus it is planned not to rely solely on the scale scores in this area.

The 13 scales scored in the final form are shown, along with the combinations which were made in Table VIII-3.

Table VIII-3

Scales Scored, SAI Final Form

Scale	Number of Items	Experimental Scale from Which Derived
1. Sociability	12	Sociability and Cheerfulness
2. Social Sensitivity	9	Social Sensitivity
3. Impulsiveness	9	Impulsiveness
4. Vigor	7	Vigor
5. Calmness	9	Calmness
6. Tidiness	11	Tidiness
7. Culture	10	Culture
8. Leadership	5	Leadership
9. Self-Confidence	12	Self-Confidence
10. Mature Personality	24	Productivity, Persistence and Responsibility
11. Conventionality	7	Conventionality
12. Theoreticality	8	Theoreticality
13. Group-Centeredness	6	Group-Centeredness
Unscored	21	All scales

Part 2 Interest Inventory

What Is An Interest?

Definitions of "interest" are sometimes made [redacted] in terms of amount of information retained in given subject matter areas (see the Information Test, Chapter VI) and in terms of overt activities (see the Student Information Blank, p.VII-[illegible]) Commonly used interest inventories define "interest" in terms of expressed liking or preference for various occupations, activities, school subjects, types of people, social situations, etc. This is the definition of "interest" intended here.

Behavioral scientists have emphasized measurement and evaluation and have analyzed interests in quite some detail. They agree that interest implies motivation--that if I am interested I will want to do something (activities), will have done something, or will have learned something (information or vocabulary).

Expressed interests are a person's opinions of himself

as contrasted with interests measured objectively by information and by reports of overt activity; interests are not verifiable or observable by an outside agent. Expressed interests may be said to be general, subjective, inclinations of the individual toward certain kinds of activities and experiences. The specific activities and experiences included in the Project Talent inventory are those familiar to high school students, i.e., occupations, school subjects, and activities.

Expressed interests, attitudes, and personality items share the characteristic that they reflect the feelings and motivations of the individual rather than demonstrate his knowledge or display his experiences. On the other hand, interests, attitudes, and personality items differ with respect to the object of the evaluation. In typical personality questionnaires, the object of evaluation is the self. Attitudes are evaluative expressions toward an institution, organization, concept, or group of people. Interests, on the other hand, deal with evaluative expressions toward occupations and activities.

Some cautions in the interpretation of interest scores are in order. Expressed interests should not be confused with ability or experience. Other factors being equal, the most useful information we can have about a person in order to tell whether or not he is likely to go to college, or is likely to achieve good grades in a course of study, is a measure of his ability and his past performance in similar school subjects. Only when these requirements are satisfied may we turn to measures of interest in order to choose between equally attractive alternatives, such as taking an

extra course in science vs. an extra course in literature.

Expressed Interests

Expressed interests have been measured in several ways. The most common way now in general use involves obtaining statements about likes and dislikes on topics such as occupations, activities, hobbies, types of people, etc. (such as the Strong Vocational Interest Blank). The other method asks the person to indicate his preference among several occupations or activities. Following is an example of the first type of question (like vs. dislike) taken from the Project Talent Interest Inventory:

> Directions: For each occupation listed below you are to consider whether or not you would like that kind of work.......Mark your answers as follows:
>
> A. I would like this very much.
> B. I would like this fairly well.
> C. Indifferent or don't know much about it.
> D. I would dislike this a little.
> E. I would dislike this very much.
>
> 1. Surgeon
> 2. Chemist
> 3. Civil engineer
> 4. Writer
> 5. Social worker

An example of the preference question, such as that used in the Kuder Preference Record follows:

> Directions: For each group of occupations listed below indicate the one you would like most, and the one you would like least:
>
> a. Mathematician
> b. Machinist
> c. Dietitian

In the first example, the student may feel that he would like all of these or none of these, and he may indicate this directly. Direct comparison can be made of the

relative strength of his interests as well as how this interest compares to others. In the second example, he must make a choice among several occupations even though he may like them equally well or dislike them to an equal degree. For the Project Talent Interest Inventory we felt that more would be gained by using the like-dislike format. It would allow us to describe our youth's interests and at the same time give us information on the number of boys and girls in the country who are interested in specific occupations such as physician, engineer, carpenter, farmer, nurse, etc.

In any case, the difference between the like-dislike format and the preference format for describing the individual's interests seems to make little practical difference. Several studies suggest that the two formats measure much the same thing, provided that the content of the questions is about the same. Thurstone ('47) used a modified paired comparison format in which the student was told to indicate whether he liked both occupations, neither one, or one more than the other. He then correlated his inventory with the Kuder Preference Record. He obtained high correlations between his scales and the similar Kuder Scales. Similar results were obtained for the Project Talent Interest Inventory and the Kuder Scales*, suggesting that the inventories are measuring much the same thing. In making predictions from inventory responses to a criterion

---

*Appendix 8b Table VIII-11

(such as successful completion of apprenticeship or training) Perry (1955) found that the preference format was better than the like-dislike format but the difference, although statistically reliable, was small.

What Are the Important Features of an Interest Inventory?

Important features of an inventory include the directions, the item format, the types of scores provided, and the content of the inventory. These characteristics of the Project Talent inventory will be described here.

The purpose of the directions to the inventory is to describe, define, and clarify the task and the context in which the questions are to be considered, and to motivate the individual to answer in a manner which truly reflects his feelings. On the Project Talent Interest Inventory the student's task is to indicate the degree to which he likes, dislikes, or is indifferent to a variety of occupations and activities. In order to clarify the context in which the questions are to be considered, he is told to consider only the activity involved in the occupations; he is not to consider associated factors such as salary, social standing, permanence of the work, training or education needed, etc. Particular emphasis is placed upon answering in terms of the degree of his liking or disliking the activity involved. This approach should help to rule out considerations extraneous to his true interests.

Item format includes the "stem", and the "options". The stem may be a word, phrase, or sentence, which describes the occupation or activity to which the student is to react. The options describe the choice of reactions, in this case a scale of like-dislike.

In the Project Talent Interest Inventory item stems were presented in the briefest, simplest form possible. We felt that statements, phrases, or complete sentences were unnecessary if the same content could be contained in a single word. Brief items would be less likely to pose a reading difficulty problem and would be likely to consume less reading time, on the average.

For each item, the student was asked to indicate the degree to which he would like or dislike the occupation or activity on the following scale:

A. I would like this very much.
B. I would like this fairly well.
C. Indifferent or don't know much about it.
D. I would dislike this a little.
E. I would dislike this very much.

Interest inventories are usually scored by adding up responses to a number of questions. Two types of scores are now commonly provided. The first, like that of the Kuder Preference Record, the Thurstone Interest Schedule, the Project Talent Interest Inventory (experimental), and others, provide scores on logically defined areas such as science, mechanics, social service, and other educational and vocationally related areas. The scales are also empirically defined. Persons scoring high on a scale usually like most of the occupations or activities in that scale; those scoring low on a scale usually dislike most of the items in the scale.

The second approach to interest scales is the occupational scale, best represented by the Strong Vocational Interest Blank, but in process of further development in the Kuder Preference Record Form D, and the Minnesota Vocational Interest Test. Occupational scales relate the

interests of the student to those of persons who have remained in an occupation for a period of years. The reasoning behind this approach is as follows: an individual is more likely to be satisfied in an occupation if his interests are similar to those of persons in that same field. Examples of the scales provided on the Strong are: Artist, Architect, Engineer, Chemist, Farmer, Social Science Teacher, etc. Clark ('56) has been developing scales for technical and skilled trades occupations with the Minnesota Vocational Interest Test. For the Project Talent Interest Inventory both types of scores will be developed for research purposes.

What Guidance Do Interest Inventory Results Provide?

Information on the prognostic value of interest inventories has been accumulated over the past fifty years through many research studies. The findings of these studies provide guidance in interpreting the educational and vocational significance of a student's interest scores. Typically, research is oriented toward some significant and measurable aspect of behavior. For example, studies of college groups relate interest scores to matriculation, choice of a college major, grade-point average, and graduation from college. Studies of occupations relate interests to choice of vocation, entry into the occupation, longevity in the occupation (but not necessarily in the specific position), success, and satisfaction with the work. Military studies have been concerned largely with selection and classification of officers and enlisted men.

How the Project Talent Interest Inventory was Developed

The Interest Inventory was included in the project to obtain from the students direct expressions of interest in occupations and activities representing a broad range of fields, sports, school subjects, hobbies, pastimes, etc.

We felt that an inventory which covered the general area of accepted instruments in present use, such as the Kuder Preference Record and the Strong Vocational Interest Blanks, should be included in order that further tests could be made of their value in occupational and educational prediction and their usefulness in guidance.

The Interest Inventory was developed by compiling a large number of items from available interest inventories and the Bureau of the Census list of occupations. The items were grouped and classified by interest areas, such as scientific, literary, clerical, etc., and then reduced to a manageable number by discarding repeated items. An experimental form was developed using 300 items out of those originally compiled. Scales for scoring the inventory were developed by assigning items to one of sixteen areas. The inventory was then administered to a group of high school and college students, some of whom had taken the Strong Vocational Interest Blank or the Kuder Preference Record. Scores from the Interest Inventory were intercorrelated among themselves and correlated with scores from the Kuder Preference Record and the Strong Vocational Interest Blank. The results show that the Interest Inventory measures somewhat the same things measured by the Kuder and Strong (Appendix 8b). The Interest Inventory was reduced to 205 items based on the results of the pilot study.

What Do We Know About Interests?

Interests, as we stated earlier, are thought to reflect an individual's motivation. These are the things we like to do, or would like to do, given the opportunity. The assumption, by behavior scientists and laymen alike, has been that measured interests will therefore relate to educational and vocational outcomes. Specifically, if our son or daughter has a high interest in scientific or mechanical pursuits, it is likely that he will do well in these courses in high school and college (provided he has the ability) and perhaps receive a degree in engineering or a related field. Beyond this series of events in the person's educational development we expect, too, that he should obtain employment in that same field, or one closely related to it; that he should remain in that field of work; and that he should achieve a measure of success and satisfaction in the field, indicated by advancement, superior income, added responsibility, or recognition for unusual or creative achievement. (Stated more succinctly, we expect interests to relate to educational and vocational outcomes, including entry, success, and completion of a related college course and entry, success, and satisfaction with the occupation.)

It behooves us to say, however, that these expectations have not yet been adequately explored. The only way to find out whether interests are important is from research studies which relate measured interests to each of these goals or criteria of performance. Some research evidence is available, enough to suggest some conclusions, but not enough for definitive conclusions. More information will be developed

over the years through research conducted as part of Project Talent.

It would not be fair to imply from our statement above that results of interest inventories should not be used, simply because exhaustive experimentation has not been done. When interpreted cautiously by a qualified guidance counselor in conjunction with information on aptitudes, achievements, plans, and personal background, interests can be useful for helping students to gain some self-insight and to help them plan their most immediate educational and vocational objectives. Final decisions for choices must rest with the student.

Much of the research work on interest inventories is largely of academic interest and will be dealt with only briefly. Some studies have been concerned with developing new inventories; some with the description and refinement of interests through the development of interest scales; some with describing concurrent relationships with aptitude, achievement, background, and personality variables. However, of most concern is what is known about how interests relate to the educational and vocational goals and criteria mentioned earlier, such as academic achievement in high school and college, successful completion of college, entry into an occupation, and success in that occupation. Studies will be discussed in terms of significance for education and for an occupation.

Significance for Education

Interests have been related to quality grades in courses, choice of a college major, and successful completion

of courses of study. Not the least of these studies are those conducted by the armed services for the purpose of selecting the most promising candidates for specialized training courses.

Green (1952) summarized the results of studies relating interests to scholastic achievement. He concluded that interest scores are not good predictors of academic achievement, but do seem to have a marked effect on whether or not a course of study is completed. His conclusions appear to be open to question.

Strong (1945) found that interests measured by his scales distinguished between dental students (141) more on the basis of completion of the course than on scholarship. Of those with high interest ratings (A or B+) 92% graduated; of those with moderately high interest ratings (B or B-) 67% graduated; of those with low ratings (C) only 25% graduated. The only difference found in scholarship was that those with the lowest rating (C) had inferior grades.

Tupes (1953) found no relationship to grades in Air Force Officer Candidate School using the Strong Vocational Interest Blank, Kuder Preference, and Guilford-Schneiderman-Zimmerman Interest Survey.

Several studies suggest that the value of interests for predicting scholarship needs to be explored by constructing special scales for this purpose. Rust and Ryan (1954) found that the Strong Vocational Interest Blank could distinguish between good and poor scholarship students after previous school record and achievement test scores had been controlled (i.e., under-achievers vs. over-achievers). This was done by constructing a new scale from the Strong

Vocational Interest Blank items. Stone ('60), too, constructed a special key from the Strong Vocational Interest Blank that predicted (r = .60) shorthand proficiency scores for an advanced junior college shorthand course. And Tussing (1942) found that a special scale constructed from the Strong predicted first semester college grades about as well as an achievement test (ACE, r = .42 vs. .39).

THE ORIGINAL DOCUMENT WAS OF POOR QUALITY. BEST POSSIBLE REPRODUCTION FROM COPY FURNISHED ASTIA

On the other hand, Super (1947) reviewed seven research reports which related the scores provided on the Kuder Preference Record and school grades. The results show a wide range of relationships (r = .00 to .60, median about .30) depending in part on the courses studied. Stronger relationships were noted for scientific than for non-scientific subjects.

These studies point up two reasons for cautious interpretation of interest scores as related to school grades. First, the scales that are available may not be fruitful for this purpose, although the construction of special scales may bring out the relationship. This suggests that basically interests do relate to scholastic achievement, but that the relationship must be brought out by combining the items in a way that differs from some of the recently accepted scale scores. Second, the importance and size of the relationship appears to depend in part on the type of course. This might mean that the amount of motivation needed varies from subject matter to subject matter. If this is the case, it would be valuable to know the student's reasons for taking each course. Is it a course required for graduation? for preparation for college? for a specific vocational objective? If the course was an elective, was it taken because the student was really

interested in it? because his parents thought it would be useful? because it appears to be an easy course? or perhaps because a teacher suggested it? If properly investigated, these reasons might be found to modify and clarify the relationship between interests and scholarship.

The question of the relationship between interests and completion of a course of study deserves additional attention. Strong ('55) found that a greater proportion of dental students with high interest ratings tended to complete training. Bolanovich and Goodman (1944) found that among women in training for electrical engineering, those with higher scientific and computational interest, and with low persuasive interest on the Kuder Preference Record tended to complete training. Berdie (1955) found that the Strong Vocational Interest Blank predicted the area in which college students received their degree. The measure was administered in the freshman year and followed up to graduation. In a more recent study (Berdie, 1960) found that college graduates from medicine, law, and accounting could be distinguished from one another from their Strong Vocational Interest Blank scores and profiles, <u>even though the questionnaires were completed in the senior year of high school</u>.

On the other hand, research conducted by the Army Air Corps during World War II on the selection of candidates for pilot and navigator training (Guilford, 1947) found no relationship between interest inventory scores and completion of training. However, further research showed that, to a degree, interests are relevant to completion of training. A few interest questions were found to aid in prediction of the criterion. Since World War II, each of the armed services

has found some interest questions to be useful in selecting and classifying officers and enlisted men for training in service specialties or for duty assignments. Clark (1956), in particular, showed that interest scales constructed specifically for the task, distinguished between men in four Navy rates. These results should again point to the need for caution in interpreting the results of interest inventory scores.

Now where does guidance and counseling fit into the picture of the direction of interests? Sarbin and Anderson (1942) found that a group of men and women who had requested occupational counseling because of dissatisfaction with their field (and who were diagnosed by the counselor as being in inappropriate fields) rarely had primary interest patterns on the Strong Vocational Interest Blank/ for the jobs they were in. And Stern, Lewis, and Bever (1958) found that vocational counseling of adults resulted in modification of interest scores. The change appeared to be toward more realistic choices.

In summary, it appears that interest/ scores relate to scholastic achievement and completion of training to at least some degree, but that new scales may have to be constructed to bring out the relationship. Interest scores may also be useful in identifying persons with inappropriate vocational goals and guiding them to more satisfying positions.

Two possible reasons that may account for failure to find relationships might be self-selection--persons who are not interested do not enroll in the course; and administrative selection--course requirements may weed out those

who are not interested.

Significance for an Occupation

If interests reveal something of the individual's motivation in his educational pursuits, they should also reveal something about his motivation in his pursuit of a career. Educational and vocational objectives presumably have elements in common in the sense that education represents preparation for some vocational objective. We expect that in addition to his education a person's interests should lead him to seek employment in a field; that he will usually stay in that field; and that he should achieve some measure of success and satisfaction in the field as indicated by advancement, superior income, greater responsibility, or recognition for outstanding achievement.

Most of what we know about interests and occupations at the present time has to do with the "group membership" and satisfaction criteria. By "group membership" we mean that an individual has entered and remained in the occupation for a number of years. Information on success as indicated by advancement, income, or responsibility is scarce, but should become available as results from Project Talent accumulate.

Results to date indicate that interests do play a part in determining what occupation is entered and whether or not the individual will remain in the occupation. Strong's research ('43) comprises the major work done in this area. His approach usually involved measuring interests of college students and then following up after various time periods to find out what occupations they had entered. His work showed that interests do indicate the occupation in which

the person is likely to be engaged. The higher the interest rating the greater the chance that the individual will be employed in the occupation (1955). But the highest rating did not necessarily indicate the occupation entered. Quite often if the occupation entered differed from the first or second choice, it still numbered among the first five choices.

To illustrate, in one study Strong (1953) followed up a group of Stanford University freshmen 19 years later (1930-1949). Of those who had expressed a choice of an occupation, 61% followed that choice. This amounted to 38% of the total group. About 50% of the total group were in an occupation closely related to their first choice. Those who changed to unrelated occupations scored lower on their first choice on the Strong Vocational Interest Blank than did those who followed their choice.

Can interests be used successfully in industrial selection? The answer appears to be in the affirmative. Bolanovich (1948) used a measure of interests in an industrial setting to select for employment women with no previous experience in factory work. By constructing a special scale for the purpose, and selecting only those among the highest 60%, the company was able to reduce a high rate of turn-over by about 50% in the first two months of employment.

Now what of interests and satisfaction with the occupation? The results seem to show that satisfaction with the occupation after a number of years is intimately related to staying in the field or leaving the field. Generally, research seems to show that about 80% of employed men are reasonably well satisfied with their occupations (Strong, 1955, Chapter 10) and that this figure varies somewhat with

the occupation and the amount of freedom in the occupation. Professional men (physicians, corporation presidents, lawyers, educators, and authors) are most often well satisfied; office workers, sales managers, and scientists are less satisfied; and skilled workmen and laborers least satisfied. Thus, for the most part, the men in an occupation after a number of years are reasonably well satisfied with the occupation. Since most of the men are satisfied, relationships between degree of interest in the occupation and satisfaction are small. That the dissatisfied tend to leave the field is shown in Gustad's (1959) study which differentiated among college professors who remained in teaching and those who left for other pursuits. Those who left were more often dissatisfied with the work and the situation. It appears, then, that interests may lead to entry into an occupation, but that dissatisfaction may result in leaving the occupation. Of those remaining, there is then relatively little difference in their satisfaction or interest.

Success in a field is perhaps even more difficult to define than satisfaction. However, relative income has some bearing on the matter. Strong (1945) showed that a life insurance salesman with A ratings on his scale earns more than do those with lower ratings. Unfortunately, no other studies which relate interest to income or any other criteria of success in an occupation are available. It may be that the relationship illustrated here would be relevant only for salesmen since salesmen are most often paid according to their production.

In summary, interests appear to have some bearing on choice of and progress in an occupation, but the evidence is

still scarce for many occupations.

There is much more to be said about interests and their importance and relevance for occupational choice and progress--too much to be included here. For greater insights into the development, meaning, and utility of interest inventories, see the references cited.

Part 3. The Student Information Blank

What Kinds of Information?

The information requested from the boys and girls in the Student Information Blank included questions on personal experiences, plans for the future, and the social-economic context provided by the family and the home. The purpose in gathering this information is to explore the significance of these events for attaining educational and occupational goals. The significance of experience and background is recognized in literature as well as science. Quotations such as, "What is past is prologue", "Just as the twig is bent, the tree's inclined", and "The child is father of the man", all recognize the molding force of our past experience. Prospective employers look at work experiences in order to decide whom to employ. Likewise, professional athletes and musicians are drawn from the ranks of amateurs. Activities and hobbies pursued in our youth often have dramatic significance for our later educational and vocational pursuits.

But what of opportunity? Many students, were they able to afford it, might like to learn how to play the saxophone, or to go to college. There is little question that opportunity to gain important experiences differs even when interest and ability are present. One of the most important indications of differences in opportunity lies with the family and home background. The social and economic provision that parents are able to give their children molds and modifies experiences available to the child.

Plans and decisions made during high school are an outgrowth of both background and experience as well as a look to the future. This information provides an account of the student's next steps toward achievement of his goals.

Some examples will illustrate questions on each topic. Among the questions on personal experiences were those exploring hobbies, sports, reading, study habits, work experiences, and the like. Among the plans and decisions about which information was requested were plans for college, for an occupation, and for marriage. Questions on the family and the home included information on the make-up of the family, parents' occupations and education, financial situation, activities, and the like.

The specific purposes for asking questions of each type also differ according to the topic. Several illustrations will clarify this point. Among the questions on personal experiences are a number which ask about hobbies, activities, sports, organizations, and the like. The reasoning behind questions of this type, generally, is to find out how the student spends his time. More specifically,

these activities reflect a student's interests and motivations evaluated in terms of what he has done in the past (rather than what he likes). Research studies have found, too, that the number and type of activities and organizations in which an individual has participated are often related to his choice of a curriculum in high school or college, his success in college, or his success in an occupation. The main point, then, is that the activities in which a student has engaged reflect a basic aspect of his make-up, and that from a knowledge of these activities it is possible to predict, to some degree, his success in educational and occupational endeavors.

Facts about the student's family and home are useful for quite another purpose. In the search for talent, research has shown that many well-qualified high school students do not enter college because their families are not able to provide them with needed financial support. For this reason questions were included on the parents' occupations, education, and income level. This information provides guide lines for determining parents' ability to pay for college.

Still another purpose applies to a student's plans for after high school. Plans are a reflection of the individual's past and his prognosis for his future. Plans may be realistic or fantasy-bound; they may be easily attained or unattainable. The skilled vocational counselor, taking into account the student's ability, scholastic achievements, interests, activities, family background, and plans, is qualified to present this information to the student, point out inconsistencies, and help him develop plans and goals

that are realistic and attainable. For those students who have not clearly formulated their plans, the vocational counselor can offer guidance toward a realistic choice.

What Kinds of Questions?

Information about background and experience of this type places major emphasis (though not exclusively) upon questions of objective fact. This emphasis lies in contrast with the Interests Inventory which asks for subjective expressions of liking or disliking. Many topics are identical with those used in Interest Inventories, such as hobbies, activities, etc. However, here the student is asked to report what he has done rather than what he likes to do.

Another way of viewing the objectivity of these questions is to say that they could be observed and verified by another person. This is not a characteristic of subjective reports. Accuracy of the report depends primarily on opportunity to observe the event, memory, and honesty in reporting.

Particular experiences varying as widely as they do, the choices (response options) provided to a question must exhaust the possibilities. This means that each individual must have an opportunity to respond in accordance with his own experience.

For any given question, each choice must also be exclusive of the other, unless instructions specifically allow marking several choices. Otherwise, confusion will arise as to which choice is the appropriate one.

Questions on plans and decisions are not objective to the same degree as are background and experiences. That is to say that they are not events that can be directly observed

by another person at the time they are reported. On the other hand, the execution of a plan of action may be directly observed. For example, if a student says that he plans to attend the state university in the coming year, the execution of this plan can be directly observed and verified. This characteristic of potential verification distinguishes plans from inventoried interests and other subjective reports, and from biographical information about past events.

In summary, questions about background and experience are objective and verifiable; questions on plans and decisions are potentially verifiable. They are distinguished from subjective reports on this basis.

How is This Information Used?

Much of the information obtained in the Student Information Blank (SIB) is to be studied for a variety of purposes. It should be helpful to the reader to understand some of the purposes for which information of this sort is gathered. Three types of studies that will be considered, representing divergent purposes, are the survey study, the developmental study, and the selection study.

The survey study is perhaps the most widely known and understood of the three types. Typical survey studies are those conducted by the U.S. Bureau of the Census, public opinion polls, and market surveys.

The primary purpose of the survey study is to determine the present status of the topic in question. For example, the U.S. Bureau of the Census gathers information on the population in the country, on persons employed in various occupations, on the age-breakdown of the population, etc. If several similar studies are conducted at different times,

it is possible to study time trends by comparing earlier figures with later figures. One such comparison is the proportion of adolescents enrolled in high school for each ten-year period for the past fifty years. Comparisons of this sort are useful for evaluating progress or decline over the years.

Developmental studies focus their attention more fully on the individual and his development physically, socially, psychologically, or economically. In a study of time trends, emphasis is placed upon numbers of people falling in a given classification; for example, the number of persons employed in skilled occupations. Emphasis is upon the classification scheme because these may or may not be the same persons so employed ten years earlier. In a developmental study, it might be found that an individual classified in the skilled occupations at one time might be classified as a foreman or manager at another time. The people are the same but the applicable classification has changed. This type of study provides better information for studying individual development and progress.

Project Talent combines features of the survey study and the developmental study. The status of the country with respect to human resources will be evaluated, but periodic follow-up studies will investigate the development of all of the students in their educational and occupational pursuits.

Selection studies focus attention on the personnel requirements of an organization. Studies of this sort may be instituted when the number of available applicants is larger than the number of positions to be filled. Tests,

questionnaires, and inventories are employed on the basis of their ability to discriminate among those with the greatest probability of success in the position. For example, during World War II the Army Air Forces instituted a program to select pilots, navigators, and bombardiers from among the many applicants for training. The result was the development of a battery of tests and questionnaires that indicated the probability of success of each applicant. This program was highly successful in that a small number of men were accepted for training, but a high proportion of those accepted graduated. Formerly, a large number of men were accepted for training, but only a small percentage graduated. The savings to the government in training costs mounted into millions of dollars, not to mention the crucial factors of time and facilities in a military operation. Selection studies, then, serve the needs of an organization. They do not give us information on the individual's development other than that available for the benefit of the organization. Information needed by an organization for predicting probable success in a position is usually much narrower than that needed to understand personal growth and development.

The purposes for which information is used in developmental and selection studies are prediction and diagnosis. Prediction is based on statistical data which show the probability of certain outcomes which correspond to given test scores or questionnaire responses. For example, it is known that the probability of attending college is high when the student's ability is high, or when his father is a professional man. The student may be apprised of his chances for entering college on the basis of the known probabilities. In a selection situation, applicants are accepted or rejected on the basis of known probabilities of success.

Diagnosis involves individual study looking toward corrective action. In practical situations this must often be done when information on probability of success is difficult to predict. In so far as possible, the skilled diagnostician uses information on probabilities when it is available. However, too often sufficient information is lacking and the diagnostician must rely on his experience, insight, and intuition for making predictions.

What is Now Known?

Present knowledge can best be evaluated by considering the significance of background, experiences, and plans for educational and occupational goals separately. Special mention will be made of the many research studies conducted by the military services for the purpose of selecting and classifying personnel for special training schools and duty assignments.

Before proceeding further, the reader should be cautioned that data now available have been gathered from many research studies conducted for a variety of purposes. Some of these studies have been conducted on a small scale to explore new research areas. Others have been conducted on a larger scale but for limited purposes. Project Talent aims to explore and integrate the significant behaviors uncovered by previous research, and to investigate their importance for development in an education or an occupation. Studies now available can provide a framework upon which to build for understanding the significance of background and experience for educational and occupational goals. However, more definitive answers to the complex set of behaviors involved in entering college, choosing a major,

achieving acceptable grades, graduating from college, entering an occupation, and progressing in the occupation cannot be answered until results from this study or others broadly based become available.

Education and Occupations

It is common knowledge that the amount of education one has relates closely to the occupation entered. This reflects both occupational requirements for college training and a series of choices and decisions oriented toward realizing an occupational goal. For example, if a young man chooses medicine as a career, the medical profession specifies that he must fulfill certain requirements, the culmination of which is graduation from a recognized medical school and completion of internship training. The student at the same time recognizes that in order to achieve this end he must graduate from high school, college, and medical school. It is the joint operation of requirements set down by some segment of society (in this example, the medical profession) and the choices and decisions of individuals who are cognizant of these requirements, that structure the close relationship of education and an occupation.

We have perhaps overstated the case in presenting an occupation with clear-cut requirements. In many business and managerial occupations and in the arts, formal college training is not so clearly related. Nonetheless, the generalization appears to be true that the amount of education one has delimits the occupation one can enter.

Studies have been conducted which describe this situation more exactly. One study (Havemann and West, 1952)

compared the occupations entered by college graduates with the occupations of /those who had not graduated from college. For men, 50% of the graduates but only 3% of the non-graduates were in professional occupations; 34% of the graduates and only 13% of the non-graduates were proprietors, managers, or executives; 10% of the graduates and 12% of the non-graduates were in clerical or sales fields; 5% of the graduates and 58% of the non-graduates were in skilled, semi-skilled, and unskilled work; and only 1% of the graduates but 14% of the non-graduates were farmers or farm workers. This study provides a good illustration of the relationship of education to an occupation. But this does not mean, of course, that the only road to success is a college education. For many persons this will be true but for others it will not. The young man with good potential as a skilled craftsman but poor potential as a college freshman, would be ill-advised to attend college simply because college graduates more often than non-graduates enter the professions or become executives. Other important characteristics (that will be discussed shortly) that are also typical of college graduates are: above average academic ability and achievement; interest and aspiration for an occupation for which a college degree is necessary; parents who want their children to have a college degree; and willingness on the part of the parents to provide financial support. With these reservations in judging the appropriateness of college for an individual youth, it is quite clear that education acquired and occupation entered are intimately related.

Significance for Education

We seek to answer questions regarding the backgrounds,

experiences, and plans which characterize those of our boys and girls who quit high school--who go on to college,--who do well academically in high school or college,--or who continue on to graduate or professional training. Identification of these students must be made in order to be able to provide them with proper guidance.

Still a problem in our nation is the large number of our youth who do not complete high school. Most states have a compulsory attendance law, requiring attendance until the 16th birthday. And it is at this point that an alarming number of boys and girls choose to forego receiving a high school diploma. A number of research workers have been concerned with this problem and have investigated the background and experiences characteristic of boys and girls who have voluntarily dropped out of high school. One study (Dresher, 1954) conducted in Michigan found that youths who dropped out of high school could be described as follows: they had failed subjects in elementary school; they were frequently absent from class; they had low scholastic aptitude; and they were often disciplinary problems. Usually they were not employed outside of school, did not participate in extra-curricular activities, and generally had no feeling of identification with the school.

Characteristics uncovered by another study (Taylor, 1955) agree essentially with those above but add the following: school drop-outs are usually older than their classmates, have less ability, are less well prepared academically, and come from a lower socio-economic background.

Having completed high school, many of our youth go on to college. The proportion in attendance appears to be

increasing continually, current estimates being in the neighborhood of 35-40%. Our next question, then, is what characteristics distinguish the youths who enter college from those who do not. Several survey studies (Berdie, 1954; ETS, 1957; Little, 1958; Stouffer, 1959) suggest the following: High school students who go on to college are above average in both ability and scholastic achievement. Economically, their families are not necessarily well-to-do, but are less often in the lowest economic bracket. Family subsistence comes before a higher education. Background and values reflected in the father's occupation play a part. Berdie (1954) notes that 90% of the students with high ability whose fathers were in top-level occupations planned to attend college. In another group with the same level of ability, but whose fathers were factory laborers, only 55% planned to attend college. Berdie notes further that the differences favored the children of the fathers with higher-level occupations even more when plans were related to actual attendance at college. Some of the underlying values associated with the father's occupation are his own education, the parents' views on the importance of a college education, and their willingness (as well as ability) to provide financial support for their children. More directly for the student, his own values play a part. Does he see college as a step toward success? Is he very interested in college? For what reasons? Social considerations play a part. Are his friends going to college? Is college social life a major reason for going to college?

Proximity to a college plays a role. Boys and girls who live within commuting distance of a college attend in

larger numbers than do those living beyond commuting distance. This affects youth living in rural areas and in regions in the West with low population density.

In summary, whether or not a student enters college is determined by a large number of variables, among which the most important appear to be his ability, his scholastic achievement, and his father's occupation and education.

Entering college does not, of course, guarantee graduation. Witness the large number of drop-outs during the freshman year. Most state universities are required by law to admit resident graduates of high schools in the state. As a result, a large number (as many as 30%) of students find the work too difficult, re-examine their goals, and drop out within the first year. Others are dropped for academic failure. Approximate figures (Wolfle, 1954) indicate that roughly 54% of those who enter college continue on to receive the degree. About one-third to one-half of those who do not graduate leave during the freshman year, the remainder being spread over the next three years.

While at college, the question of scholarship is important to both the individual and the university. Private colleges and universities often select their students on the basis of academic performance in high school and objective tests of academic achievement. State universities have frequently been prevented by law from selecting among applicants within the state, but have used tests and high school academic standing to select among out-of-state applicants. Both variables are quite often used for advising students on their choice of a major course of study. Whether for selection or for individual guidance, tests of

academic achievement and records of academic performance in high school are among the major indicators of academic potential in college. Again, however, other characteristics come into play. The most consistent among these other variables appears to be a set of motivational characteristics reflected in the student's habits and attitudes toward studying. Habits and attitudes toward studying are reflected in questions such as the following taken from the SIB:

> "65. I do a little more than the course requires."
>
> "66. I make sure I understand what I am to do before I start an assignment."

Some recent research ~~[illegible]~~ appears to show that these qualities of study methods, effort, and attitudes contribute information on probable achievement in college beyond that provided by tests of academic achievement (Brown and Holtzman, 1955; ~~[illegible]~~ 1956; Ward, 1959). This information should be considered with reservation since the number of studies which have investigated the importance of study habits is still small and results are inconclusive.

A host of additional background and experience characteristics have been investigated in order to understand what, other than ability and past achievement, contributes to good scholarship in college. Surprisingly, once in college the education and occupation of the parents make little difference in how well the student does in his courses. Other characteristics that have been investigated in limited settings are: participation in high school

activities; participation in out-of-school organizations; offices held in organizations; size of the community in which they reside; order of birth among their siblings; being an only child; etc. Results are as yet too inconclusive to afford any generalizations concerning this material. Probably of much greater importance than any of these background and experience characteristics will be the demands and competition afforded by the university the student attends and the program of study which he pursues. Different universities, and even specific schools within universities, have varying standards for scholastic achievement. Colleges with reputations for high standards tend to draw more applications from academically qualified students, on the average. Engineering and science programs are usually more demanding than many other programs in the same university. Complete understanding of why students of equal ability achieve different quality grades will certainly have to be interpreted in the light of the university he is attending and his program of study. Each may exact different motivational demands and draw students of varying backgrounds and experiences.

For the majority of college students, receiving the Bachelor's degree is the final step in their formal education. Others choose to continue for a post-graduate or professional degree. Who are these people? What are they like? What are their aims and goals in seeking advanced training? From what kinds of backgrounds do they come? These are some of the questions that must be investigated if we are to learn how to guide toward advanced college training those of our youth who are interested and able to master the work. Unfortunately, at this time questions are

more plentiful than answers.

As with college training generally, the number of students who go on to graduate or professional training is increasing rapidly over the years. For example, the number of Doctor's degrees conferred in 1911-15 was about 2,500; in 1931-35 about 13,500; and in 1946-50 about 21,000 (Wolfle, 1954, pp. 300-301). These figures give some indication of the supply and demand for advanced specialized training at the college level. The figures do not, of course, indicate how many students enter advanced training and terminate their education with a Master's degree. Nor do they indicate how many enroll and subsequently drop out without receiving the degree. We need information that will help a student appraise his probable success or failure in graduate or professional school before he enrolls. Those students with little chance of success could be so advised, while those who appear to have the necessary qualifications could be encouraged to attend.

What has research uncovered so far? Ability level makes a difference, but not as much as might be expected (Wolfle, 1943, pp. 319-322). Grade achievement while in college makes a greater difference (Gropper, 1959; Havemann and West, 1952), as does the student's purpose in seeking advanced training. The purpose is usually to gain greater specialization in a chosen vocation. One study (Gropper, 1959) found background differences between graduate school students and candidates for professional degrees. Those attending professional schools tend to make their decisions for advanced training at an earlier time than do those in graduate school. They appear to be more strongly influenced

by parents and other relatives, and more often have fathers who have professional degrees. They tend to receive financial support from their parents rather than from other sources. Graduate students tend to make their decisions to attend graduate school at a later time, are not often influenced by parents or other relatives, and usually obtain financial support through assistantships and fellowships. They often have fathers who have also had graduate work. About 28% of both groups entered graduate study after being away from college for several years.

This appears to be as much as we know about the backgrounds of our graduate and professional students. What we need is information on the habits of study, reading, and value for education in early life as well as during college. From common observation we would expect graduate and professional students to display greater active interest in scholarship and erudition for their own sake. We would expect their backgrounds to encourage and value an education. These characteristics have not yet been fully explored.

Significance for an Occupation

Having completed one's education, attention focuses on occupational endeavors. What occupation shall I choose? Where shall I apply for a position? Where am I likely to achieve some measure of success and satisfaction? What am I qualified to do with my capabilities and education? For some, the answer is quite clear. They have obtained college or vocational training that qualified them for a specific occupation. Others, with less specialized education and training must still choose from a number of possible positions. What occupations do our young people enter?

How successful will they be in their occupational pursuits? What backgrounds and experiences lead to entry or success in one or another occupation? These are the questions that must be answered in order to be able to guide our youth toward appropriate occupational goals.

What we mean by "entry" and "success" in an occupation will help to clarify the material that is to follow. Entry means simply that the individual has applied and has been accepted for a position in some occupation. He may have entered the occupation as a result of careful planning and by conscious choice; or he may have entered the occupation simply because of immediately available opportunities, having been encouraged to apply for the position by a friend or relative, or simply to support himself. Of course people enter and leave a number of positions during the course of their careers; however, there seem to be relatively few who move from one major occupational group to another (i.e., from professional to managerial or sales, etc.) very frequently. This criterion of entry or <u>occupational group membership</u>, as it has been called, provides a good picture of the kinds of people who are in different occupations for a specified time period.

Success is a bit more difficult to define but can be attacked sensibly if we specify what we mean by success in an occupation. Success may be defined objectively and subjectively. One way of defining success objectively is in terms of the occupation itself. Thus, physicians, scientists, writers, and most professional persons may be regarded as being at least moderately successful by virtue of the fact that they occupy a position high on the occupational ladder

in a competitive situation. This is also true for many highly skilled occupations. Obviously, this is the same thing that we referred to as entry or occupational group membership.

A second way of defining success is in terms of income. Presumably, a man's worth in an occupation is reflected by the amount he earns. Certainly, many other factors play a part in earnings; however, success should be at least in part reflected in earnings. Current evidence seems to show that this is true in some occupations but not necessarily in others.

Other ways of defining success are specific to some occupations. These include the attainment of supervisory responsibility, piece-work production, articles and books written, or receipt of awards or recognition for achievement.

Subjective definitions include how successful a man feels he has been compared to his fellow-workers, and how successful he has been compared to his own standards and aspirations. The former is a "group-oriented" standard while the latter is a personal standard. Some work has been done using the "group-oriented" standard or self-rating, but to this writer's knowledge, no work has been done which relates ability and background characteristics to the personal standard. That the two are different appears to be quite clear. That is, a person may feel that he has done well compared to his fellows, but may not feel satisfied that he has achieved his full potential.

The well-known merit-rating is a common way of defining success, particularly for those positions for which no

concrete form of production or output is available. In this case the supervisor or superior evaluates the individual on the quality of his work. Ideally, if the evaluation is to be a fair one, the supervisor should have had equal opportunity to observe all of those under him; he should approach his task analytically and objectively; and he should focus his attention on those aspects of the work that are critical.

Generally, there appear to be three contexts in which studies of occupations have been made. The Bureau of the Census provides information each decade on the number of persons employed in various occupations. Although useful for their purpose, this material does not very often relate occupations to background characteristics of individuals. Those background characteristics that have been related are sex, age, and education. The only "success" characteristic that has been related to occupation is annual income. It should be quite clear from one's own observation that these characteristics are highly interrelated.

A second group of studies are those conducted by individual companies, organizations, or industries for the purpose of selecting the most promising employment applicants. In this type of study, personal characteristics are related to specific aspects of the work indicative of "success". This might include any or several of the success characteristics mentioned earlier, including supervisor's ratings, production records, or amount of time with the company.

A third group might best be called "career-pattern" studies. This type emphasizes the individual's progress in an occupation. He may change positions or occupations, may

be promoted, handle greater responsibilities, etc. Whatever his role or course of action, the emphasis is on individual choice, action, and progress. Another way of viewing this type of study is from the point of view of human development. The individual's progress is followed through education, an occupation, and retirement, documenting the pattern of his life history as it progresses.

The Bureau of the Census publishes figures on the relationship of occupations to age, sex, education, and income. In general, these data show that as people grow older they progress to higher positions and their earnings increase. Those of us with more education, on the average, occupy higher positions and receive more income than those with less education. Focusing our attention on sex alone, we find that women obtain less education than men, are concentrated in different occupations (teaching, nursing, secretarial and clerical work, etc.) from men, and receive less income, on the average. Of course, many women become housewives and are not included in the labor market. The fact that they may be well-trained stenographers, nurses, teachers, or even scientists is not shown in this information.

The Census information is useful to a limited extent for understanding occupational growth, but it leaves many questions unanswered. For example, how are people guided into different occupations? What relationship exists between the father's occupation and the occupation entered by the son or daughter? Or do American parents stress bettering one's position to such an extent that the child's occupation is higher on the occupational ladder than was

the father's? How is the economic bracket of the parents related to the occupation and economic bracket of the children? Parents in higher income brackets are likely to be in a better position to educate their children, provide them with a financial base that sometimes extends into the married years, provide aid for business ventures, introduce them to friends who may provide help in obtaining a desirable position, etc. Furthermore, the children usually inherit the parental estate, thus bettering their own financial position. Many questions of this type will have to be answered before a comprehensive picture can be obtained of the variables that enter into choice of an occupation and resultant income bracket.

Havemann and West (1952) limited their study to college graduates. They too found that age and sex make a difference in occupation and income, but they went on to other findings. They found that married men earn more money than do single men, and after age 30 married men with two or more children earn more than do those with no children or with one child. There may be several reasons for these findings. The first that might occur to most of us is that married men, particularly those with several children, have a greater incentive to work hard to provide a living for their families. A different answer might lie in the personality and other characteristics of men who marry and those who do not. Perhaps both marriage and financial success may reflect the sociability, drive, and ambition of the person. This we will learn later when we use the married-not married criterion for "validation" type studies. Still another possibility lies in pay and

promotion considerations carried out by employers. Married men with children may be given greater consideration for advancement than those with no children and single men.

Another approach was taken by Thorndike and Hagen (1959). They followed up about 10,000 men who had taken the Army Air Corps selection tests (for pilot, navigator, and bombardier training) during World War II. The tests included measures of aptitude and ability specifically designed for the purpose of selecting Aviation Cadets, but also included a large number of background questions including (1) family and personal background; (2) major subject in college, if they attended; (3) success with different school subjects; (4) participation and skill in sports; (5) other activities; (6) hobbies; (7) work experiences; and (8) reasons for seeking Aviation Cadet training. Comparisons were made between groups of men who had entered a given occupation with the entire group that had been tested. Comparisons of this sort reveal the extent to which accountants, for example, differ in background from men in all other occupations. In all, over 100 questions were used to portray differences among the men in the occupations studied, making the detailed results difficult to report. Suffice it to say that almost all of the differences made sense. As we would expect, college attendance and college major accounted for the largest number of differences among the occupational groups. This is simply a reflection of educational requirements for most occupations. Beyond these results, the differences are more revealing. In most cases the background characteristics appeared to be quite relevant for the occupation entered. Some examples will

clarify the point. Skilled tradesmen showed a history of relatively poor academic school work, but took shop and mechanical drawing courses while in school. Hobbies and free-time activities were often mechanical in nature. Previous work experience before entering the Air Force was in skilled trades. Men in professional occupations usually did well in school, particularly in intellectual courses. Hobbies and activities were not often of a mechanical nature.

These authors tried relating test scores and background characteristics to income for the men in each occupation, but without success. The only differences in income that were found could be accounted for by differences in the occupations themselves. This means that average income differed between accountants, engineers, and teachers, but that the tests and background characteristics did not distinguish among those accountants with higher or lower income than the average.

Studies by individual companies and organizations usually have the purpose of selecting from among a number of job applicants. Tests and personal background information often indicate the most promising among a large number of candidates. In this setting, much of the background information obtained is requested on the application form. Responses to questions are scored or weighted, giving the often used title, "Weighted Application Blank".

The first use of weighted and keyed application blanks was by the Life Insurance Agency Management Association. Kelly (1942) developed a Biographical Inventory for the Committee on Aviation Psychology of the National Research Council using funds supplied by the Civil Aeronautics Administration. This inventory is the basis for special forms developed by the

Armed Services for selection and assignment of men to specialty schools and duty assignments. Studies conducted by the Armed Services will be discussed in the next section.

Although concerned with relative degrees of success among a group of job applicants, this type of work places less emphasis on the individual and more on satisfying the requirements of the organization. Fortunately, the needs of the individual and the requirements of the company are not often in conflict. However, the point to keep in mind is that the weighting of the application blank for background information is based on the company's past experience on criteria such as production, sales, supervisory ratings, length of service, etc. These are all what might be termed "organization specific" criteria in that they may be appropriate to one company but not necessarily to another.

The information obtained usually parallels the material outlined for the Thorndike and Hagen (1959) study. In many cases, greater emphasis is placed on previous work history and previous salary, and additional questions are asked on such topics as proximity to the company, whether any friends or relatives work for the company, and the reasons for applying to this company. Moderate success has been achieved for applicant selection.

Career pattern studies focus attention on the individual's choices as he progresses through his education and occupation. Emphasis is upon individual choice and decision, whether rational or irrational, whether well- or ill-conceived. These studies attempt to account for the variables involved in choice of an occupation. Questions such as the following are posed: When do children tend to

choose an occupation realistically? Do children tend to choose their father's occupation? What influences decisions to enter an occupation--or to leave the occupation once having entered? Several examples will illustrate the type of results that have been obtained.

Vocational interests and choices by high school students below the 11th grade have been found to be unrealistic (Green, 1952, Ch. 20). Earliest choices quite often include a preponderance of "adventurous" occupations such as airplane pilot, explorer, foreign correspondent, etc. In high school, choices include more "prestige" occupations--unfortunately, more choices than available positions. By and large, present results indicate that choices are modified as the individual gains more experience and knowledge about the occupation. From the point of view of guidance and education, the problem is to minimize the number of poor choices that are made so that students do not waste time getting an education or working in an occupation or an area ill-suited to their capabilities.

Do children tend to choose their father's occupation? To some degree. But several tendencies have been noted (Jensen and Kirchner, 1955). First, exactly the same occupation is frequently chosen (or entered) when the father is a professional, proprietor, manager, clerical worker, craftsman, or operative. Others enter jobs that are at a higher level than that of their father, particularly the sons of service workers become operatives, and sons of unskilled fathers enter operative and crafts work. Needless to say, this information indicates that parents often influence their children's choice of an occupation. However,

it should be apparent, too, that the school, friends, and the labor market situation all play a role in the occupation entered.

Significance for the Military Services

Although the uses of background information by the military services might be subsumed under occupational significance, the extensive developmental work by the military groups deserves special mention. The purpose for which background information has been used in the military is to identify those individuals who would be best qualified for specialty training schools and for various duty assignments.

Historically, this developmental work was extremely important in exploring the utility of background information for selection and classification for occupational duties. During World War II, the Army Air Corps experimented with many kinds of information including personality questionnaires, interest blanks, and background information. In almost every case, the background information proved to be a valid predictor of success in pre-flight school; the personality and interest measures did not.

The types of information used were described on page___, and included general family background, work experiences, hobbies, activities, organizations, etc. The final product of their efforts was to produce scales for the pilot and navigator positions. Men scoring high on the pilot scale, for example, were likely to succeed in pilot training.

Since World War II, scales have been developed by each of the military services for technical, clerical, and administrative positions.

Project Talent's Student Information Blank

The Student Information Blank (SIB) contains questions on personal experiences, family and home background, and plans for the future. The questions are largely objective and verifiable facts or events rather than subjective reports of personal feelings. Students answered 394 questions about themselves on pre-coded answer sheets. The pre-coded answer sheet was used because it facilitates analysis on electronic data processing equipment. (A pre-coded answer sheet presents all possible responses to the individual, requiring only that he mark the choice most appropriate to his own particular situation. He is not required to write his answer.) The number of possible responses to a question ranged from 2 to 36; the majority of the questions had 6 responses.

Several examples of questions in the SIB are as follows:

A question on personal experience

"13. How many times in the last 3 years have you been captain of an athletic team?

A. None
B. Once
C. Twice
D. Three times
E. Four times
F. Five or more times"

A question on the family

"176. How many books are in your home?

A. None, or very few (0-10)
B. A few books (11-25)
C. One bookcase full (27-100)
D. Two bookcases full (101-250)
E. Three or four bookcases full (251-500)
F. A room full--a library (501 or more)"

A question on plans

"304. What is the greatest amount of education you expect to have during your life?

A. I don't expect to finish high school.
B. I expect to graduate from high school.
C. I expect to obtain vocational, business school, or junior college training.
D. I expect to obtain some (less than 4 years) regular college training.
E. I expect to graduate from a regular four-year college.
F. I expect to study for advanced college degrees."

What follows is a detailed description of the questions that were included in the SIB and the reasons for asking these questions. Because the questions are of a personal nature, every step has been taken to assure anonymity. Under no circumstances will the name of any student be disclosed. The sole purpose in asking these questions is to study the patterns of experience and background that contribute to avocational and vocational growth. The questions are grouped under personal experiences, family and the home, and plans for the future.

A. Personal experiences

1. Organizations. Questions were asked on the number and kinds of organizations in which youths have participated, the extent of their activity in the organizations, and the number of offices held. Organizational membership is one form of social participation and holding an office indicates some degree of leadership experience. This experience has been found to be related to entry and success in the military. In high school, it is indicative that the student will not be likely to drop out.

2. Hobbies and recreational activities. Questions were asked on the number and variety of activities in

which they have participated regularly, including woodworking, collecting rocks or coins, hunting or fishing, etc. These are voluntary activities indicative of an individual's interests, and have been found to be predictive of entry into a field of study and into an occupational field.

3. Work experiences. Questions included the types of work they have done, the amount of time they have worked during the school year, summer jobs they have held, and chores they do at home. These experiences are expected to relate to a student's vocational growth and maturity, and his choice of an occupation.

4. Sources of personal income. Questions were asked on the percent of the student's spending money he received from his family and from a job. The questions may indicate personal initiative and independence.

5. Dating and social activities. Questions included how old they were on their first date, when they learned to dance, and how often they go out in the evening. These questions assess social experiences important in the life of every adolescent. Social development has also been found to be related to vocational choice.

6. Reading. Questions were asked about the number of books they read, and the types of reading they do (for example, adventure, science fiction, classics, biography, etc.). Amount and type of material read is often indicative of interests that relate to educational and vocational choices.

7. Studying. Questions included the number of hours per week they spend studying and a number of questions on their approach to studying. The latter included attentiveness, reading and concentration habits, effort, orderly habits,

etc. Questions of this type have been helpful in identifying students whose habits of study are either poor or outstanding. Research has shown that study habits are predictive of scholastic achievement even when ability has been equated.

8. Schooling. Questions were asked on the curriculum in which they were enrolled, the number of courses taken in various areas (science, mathematics, English, foreign languages, etc.), quality grades in these courses, transfer from school to school, and amount of time missed from school. These questions assess preparation and scholastic achievement, and relate to entry into college and scholastic achievement in college.

9. Guidance and counseling experiences. Questions were aimed at determining what persons (counselors, parents, teachers, etc.) students approach for guidance and what problems they discuss. An indication may be obtained of the extent to which school guidance facilities are being used. Students were asked to indicate the number of times they have discussed their school work, educational and vocational plans, and personal problems with counselors, teachers, principal, parents, siblings, and friends.

10. Awards. Questions were directed at awards won in academic, artistic, athletic, and organizational endeavors. It is expected that awards will have some significance for the students' educational and vocational development.

11. Driving. Questions were directed at determining whether or not they drive a car, how old they were when they learned, whether they learned at school, and how often they have use of a car. This is a much-discussed topic for teen-

agers. Using the family (or one's own) car may be for dating, work, or a mechanical repair hobby. Therefore, the significance of driving should be evaluated in conjunction with these other activities.

12. Health. A large number of questions were asked about the students' health. The purpose of most of these questions was to identify students whose health might limit their educational and vocational opportunities. Other questions were asked to obtain a general statement of health status.

B. The Family and the Home

1. Family composition. This group of questions documents the family structure: mother, father, siblings, grandparents, and other relatives. Also included were questions on the parents' ages, and whether the student is a twin or triplet. These questions were included largely to introduce other questions on the parents' occupations and education, and the education of the student's brothers and sisters.

2. Parents' occupations. Questions were included on the parents' occupations posed in terms of broad categories such as skilled, managerial, professional, farm, etc. This occupational classification is a modification of that used by the U.S. Bureau of the Census. Other questions asked about the parents' employers, number of jobs their father holds, level of responsibility, and questions on their mother's work. The purpose of these questions is to determine the socio-economic status of the family. Parental occupation has been found to relate to the children's achievement of their educational and vocational goals.

3. Parents' education. The amount and type of the parents' education relates to the amount and type of education the student is likely to seek.

4. Economic situation. The purpose of these questions was to establish a rough index of the ability of the family to support the student while in college. This will help to determine the need for scholarships. The questions included an estimate of the total family income (in broad categories) and the major source of income (wages, professional fees, etc.).

5. Description of the home. The purpose of these questions was to obtain supplementary socio-economic indicators. Questions were directed at whether they are apartment or home dwellers, the amount of rent paid or the value of the home; number of people in the home; and the number and type of articles, appliances, tools, and furnishings.

6. Family mobility. The purpose of these questions was to determine the significance for our youth of family migration--intra-regional, city-to-suburbs, house-to-house; included were questions on the number of years they have lived in their community, number of house-to-house moves, city-to-suburb moves, region of the country in which they formerly lived, etc.

7. Foreign languages spoken by parents. These questions were asked to determine the influence of parents who speak a foreign language fluently on the student's achievement in foreign languages. The present international situation has emphasized the need for more persons with specialties in foreign languages. These questions should help to identify students with good potential for learning foreign

languages.

8. Books and magazines in the home. These questions were indicative of the interest of the family in reading and the opportunity afforded the student to gain an interest in reading. Questions were directed at the number of books in the home and the number and types of magazines (news magazines, movie magazines, parents' magazines, etc.).

9. Automobiles owned. The number of automobiles owned by anyone in the home provides some indication of opportunity for the student to learn to drive and learn to repair a car. It also gives some indication of economic status even though unrelated to occupation, education, and income.

C. Plans for the Future

1. Educational plans. Questions were included on plans to quit high school; to attend college (full time or part time; junior college vs. four-year college), to attend a trade, vocational, or business school after high school graduation, and the total amount of education (high school, college, or graduate or professional) our youth expect to obtain. They were also provided an opportunity to express their reasons for or against going to college.

2. Specific college plans. This group of questions was asked only of 11th and 12th grade students. For those who expected to attend college, questions were asked on their choice of a college major, the type of college they expected to attend, the number of colleges to which they had applied, distance to college, plans to obtain a loan, the source of other financial support for college, etc. These questions were included to survey the present state

of specific college needs and the need for financial support.

3. Plans for marriage and children. The age at which the student plans to marry often relates to his plans for college. Those planning an early marriage usually have not planned to attend college. In recent years there appears to be some reversal in this trend.

4. Economic aspirations. This group of questions is directed at the student's concern for economic gain. Questions included aspirations for income, insurance, savings, securities, and real estate. It is expected that these questions may relate to economic gain in the future.

5. Plans for fulfilling military service requirements. Military service constitutes an obligation for most boys that interrupts long-range educational and vocational goals. For others, the military is a potential career. Questions included plans for entering the military, choice of service branch, and plans for a military career. This will permit estimation of what part of the national talent pool each service might expect to be available to it under various circumstances.

6. Occupational choices, decisions, and values. Questions included the specific occupation chosen; the occupation preferred; the grade in which the occupation was chosen; the number of different occupations considered; and the considerations, such as income, security, interest in the work, etc., that are important to them. These questions have proved useful for predicting the occupation entered, for evaluating maturity of choice, and for providing guidance to students on their vocational choices.

What Will Be Done With This Information?

It should be apparent to the reader that this information will serve a wide variety of purposes. First, information on experiences, background, and plans will be cross-tabulated with test scores on aptitudes and achievements. When analyzed, this information will be valuable for evaluating the nation's resources of talented youth.

Our youth will also be asked to complete questionnaires one, five, ten, and twenty years after graduation. In this way it will be possible to follow individual growth patterns, particularly those of educational and vocational significance. For example, it will be possible to study the patterns of experiences, backgrounds, plans, interests, abilities, and achievements that contribute to becoming a physician, scientist, mechanic, nurse, stenographer, and so forth. It is to the better identification and development of the talents of our youth that we address our efforts.

\- - -

APPENDIX 8a

SAI Scales

Each of the scales in the tryout version is described and briefly discussed below.

1. Leadership

This trait concerns activities such as taking charge, giving orders and actively seeking responsibilities. It is commonly identified as part of the broad factor of dominance.

Sample Item:

I like to make decisions.

The tendency to seek responsibility and to seek to "take charge" is considered desirable by many people, and is a major factor in vocational and educational choices.

2. Vigor

This trait concerns the activity level of a person, primarily on the physical side. It is related to the somewhat more broad factor of General Drive which would include high levels of both physical and mental activity. It is fairly stable and easily observable in the high school student.

Sample Item:

I play games for hours without getting tired.

Measurement of this factor will make it possible to study the extent to which persons with ability and interests are successful in proportion to their drive level. It will aid in defining creativity with respect to amount of pure activity involved. It will make possible the correlation of vigor with other aspects of ability such as productivity, scholastic aptitude, achievement in various subject matter areas, overall adjustment and ingenuity.

3. <u>Productivity</u>:

This trait concerns the ability to get things done, particularly in good time. The emphasis is on quantity. No assumptions are made as to the quality of the products other than that they are acceptable, i.e., meet the required minimum standards. Essential characteristics seem to include clear understanding of task goals; strong need for achievement, oriented toward accomplishments and completed products; interests stable enough to allow the completion of products; and habitually good concentration and efficient utilization of time and effort. This trait is closely related to persistence, vigor, and responsibility.

Sample Item:

I work fast and get a lot done.

In measuring the extent to which the person tends to produce and to produce efficiently in good time, we are getting a measure of one of the aspects of human ability that society values most. These measurements form a good criterion against which to relate other kinds of personality and training and background variable. Especially interesting are relationships with bibliographic information and with school educational practices.

4. Tidiness

This involves a tendency to want to keep things neat and orderly. It is probably related to a more general factor of compulsiveness, but as a trait can be clearly defined and easily observed at the high school level.

Sample Item:

I do my homework as neatly as possible.

Tidiness is probably of lesser consequence to the study *per se*; however, to the extent that tidiness relates to rigidity or lack of adaptability on the one hand or to executive administrative or organizing ability on the other; it can be of considerable importance to prediction of vocational career, types of educational training and later successes.

5. Sociability

This refers to a tendency to like and need to be with people *i.e.* to gregariousness. It is related to the commonly found broad factor of sociability. This is a highly important and observable aspect of behavior at the high school level.

Sample Item:

I take a big part in social activities.

The importance of sociability to many of the factors being studied in the project is obvious. Few if any people really work alone and the ability and extent to which people relate to other people is a prominent and determing factor in the type and kind and success of their later vocational and educational efforts. Certainly one of the objectives of guidance and counseling has often been to assist the student in getting along with people in more adequate fashion.

6. Self-Confidence

This includes a basic ego-security manifested in confidence in one's personal worth and social acceptability. It implies willingness to proceed on one's own and a certain independence of thought and action. Factorially, it is related to the commonly found factor, self-sufficiency, though it correlates moderately well with mature personality as well. It appears important and ratable at the high school level.

Sample Item:

I'm equal to any occasion.

By and large, self-confidence is a highly desired trait. Again, measures of self-confidence can be related to many of the objectives of the study, such as the effects of guidance and counseling, relationship to bibliographic information creativity, productivity, ingenuity, achievement and the prediction of vocation and education success.

7. Cheerfulness

This ~~trait~~ concerns the tendency to "make the best of things," be lighthearted, look on the bright side, and avoid a pessimistic view. It does not relate closely to any of the six broad factors specified earlier, yet is commonly mentioned in the literature as a well-defined, fairly stable trait. It is one easily observed at the high school level, and probably of considerable importance to future behavior.

Sample Item:

I am good-natured most of the time.

This trait, a generally desirable one, probably has its most important relationship to the study with respect to the choices, plans and decisions made by high school students. These quite probably are highly influenced by general outlook on life, that is optimism, hopefulness, eagerness, cheerfulness, versus pessimism, sadness, and seriousness. It is possible that measurement on this trait will prove to be discriminating between various kinds of occupations such as service occupation as opposed to research occupations.

8. Culture

This trait concerns a tendence to appreciate and display refinement, culture, good taste, and aesthetic things.

Historically these items have been closely related to the often discussed masculinity-feminity factor. However, the culture scale seems to be much less a heterogeneous mass of correlated attributes than does the traditional masculinity-feminity factor. It is probably a little less observable in the high school setting than some traits, but bears important implications for future behavior.

Sample Item:

I enjoy works of art.

This particular trait appears to have somewhat less implication for the study as a whole, but has some bearing on the selection of hobbies, avocations and vocational choices.

9. Talkativeness

This is a clearly recognizable aspect of behavior, of considerable import. It is characterized by desire to talk about anything and everything with almost anybody, sometimes to the eventual discomfiture of the talker. Factorially, talkativeness seems to be related to vigor, to sociability and perhaps to lack of self-discipline. Even so it seems to have some essence of its own. Bright children are often talkative, as might be expected from their superior verbal facility. Yet average to dull children may also display this trait.

Sample Item:

I start conversations easily with strangers.

The measurement of talkativeness relates somewhat less well to the objectives of the overall study than other traits. Certainly, however, the relationship of talkativeness to such factors as vigor, sociability, and self-confidence, etc. would be of considerable interest, though somewhat outside

the stated measurement goals of the study. Talkativeness might also be interestingly, perhaps negatively, related to productivity, ingenuity and vocational successes.

10. Impulsiveness:

This trait concerns the tendency to make snap decisions, to act without full consideration, to do and say things on impulse and whim. It is probably related to lack of self-discipline, and perhaps indirectly to drive. This appears to be a trait which is observable, but perhaps somewhat difficult to self-rate.

Sample Item:

I usually act on the first plan that comes to mind.

The measurement of impulsiveness bears an important relationship to items such as stability of occupational and educational choice, the number of switches in program and occupation, effectiveness of counseling and factors involved in making decisions and plans.

11. Persistence

This trait refers to a tendency to keep on working on something until it is finished or accomplished. It is undoubtedly related to drive and also to self-discipline, since it is effectively drive directed toward some goal. This is observable at the high school level and is very important to future behavior.

Sample Item:

I usually stick to the things I start until I finish them.

As this trait refers to a tendency to continue and persevere in working towards some long-range goal, it is of importance to relate it to any of the long-term aspects of

the data being gathered. It is especially important in evaluating satisfaction with career choices, educational choices, and in examining the current choices against follow-up activities.

12. Calmness:

This concerns the ability to react appropriately to emotional situations rather than displaying extremes of elation, temperament, excitability, depression, etc. Evenness and smoothness of temperament are the observable characteristics and the trait is considered highly important. This trait is clearly related to the broader factor mature personality.

Sample Item:

I rarely lose my temper.

Many of the objectives of the study have to do with the extent to which the student is achieving a satisfactory adjustment with his environment and its demands. Therefore, the measurement of calmness provides a measure which is highly related and highly important to the child's ability to integrate himself into his school and work situation in such a way that he can take advantage of his abilities and prior achievements. Measurement on this trait will make it possible to study the extent to which emotional stability is related to occupational-vocational success, effectiveness of counseling, efficient decision making, etc.

13. Social Adjustment

This trait concerns reactions and interactions with others. On the one end, it concerns helpfulness and cooperation; on the other, behavior which is distinctly and overtly anti-social and/or hostile. This trait is extremely important

in the high school setting. Factorially, it is probably related to sociability, but has often been identified as a separate factor.

Sample Item:

People seem to think I'm cooperative.

Measurement along this dimension is very important in attempting to determine what conditions of background and interest and aptitude go along with the desire to help and cooperate and work efficiently with others. There is undoubtedly a great deal of overlap between this trait and sociability.

14. Theoreticality

This concerns the tendency to be a thinker rather than a doer, to prefer intellectual activity of all kinds to working with tangible things. It has been identified as a factor in some studies, but is probably related to sociability and self sufficiency. It is important to behavior and probably easily observed and rated. It is very likely to be a stable characteristic at the high school level.

Sample Item:

I spend a lot of time by myself thinking.

This dimension is extremely important with respect to the kind of vocation selected and the extent to which the person pays attention to and is concerned with society and people and things, as opposed to ideas, concepts, philosophies. It is a variable to which effectiveness of counseling and types and kinds of educational and vocational choice and decision making should be related.

15. Responsibility

This trait involves not the active seeking of responsibility

but the willingness to accept and discharge responsibilities, even though they be distasteful, to the best of one's abilities. This person has a strong sense of duty and can be depended upon to carry out his assignment even at personal discomfort. This trait is observable in high school and is factorially related to mature personality and to conformity and persistence.

<u>Sample Item</u>:

People say they can count on me.

Responsibility, willingness to accept and discharge responsibilities, is certainly one of the most important of the work attitudes or traits, and as such can reasonably be related to almost any of the variables under study.

16. <u>Social Sensitivity</u>

This trait involves the ability to put oneself in another's place. This person is aware of and concerned about the feelings and desires of others. The trait is perhaps somewhat rare in high school, but constitutes an important part of behavior. It is factorially related somewhat to sociability and mature personality.

<u>Sample Item</u>:

I don't like to see someone's feelings hurt.

17. <u>Group-centeredness</u>

This trait concerns preoccupation with the good or welfare of the group as opposed to self-centeredness. While amenable to self-rating and undoubtedly important to future behavior, it is perhaps less observable, behaviorally, at the high school level than some other traits. Factorially, it tends to be related to sociability, mature personality, and conformity.

Sample Item:

> I'd give up my place on the team, if it meant the team would win.

The trait of placing one's own advancement ahead of that of the group or vice-versa is highly important to getting along in today's complex, highly-peopled society. Measurements here might be related to many background factors, to many other personality variables, such as leadership and vigor and to various family and cultural factors, as well as vocational and educational successes.

18. Conventionalism

This concerns the trait of conforming and adapting to rules and conventions, whether liking them or not. It involves a high respect for the rule-making body (not necessarily society, but perhaps a sub-culture) with which an identification is achieved. It is related to group-orientation and sociability and negatively to self-sufficiency.

Sample Item:

> I obey rules whether I like them or not.

Perhaps the most interesting relationship would be conventionalism and the creativity test. Productivity and originality are some of the other areas where relationships with conventionalism should be of value for study. The effect of conventionalism on decision-making, including its promptness and its adequacy, and on predictability of occupation and occupational success are certainly of interest.

Pre-testing of the Student Activities Inventory

While Project personnel felt considerable confidence in the psychometric and psychological procedures which had produced the experimental form of the Student Activities Inventory, it was nevertheless felt to be essential that the experimental form of the test be pre-tested on a substantial number of students. In this way, items and scales could be refined and administrative problems taken care of. Therefore, arrangements were made to give the Inventory to several hundred students in the Fall of 1959.

A. The Administration of the Experimental Test, Fall 1959

The experimental form of the test was given to several groups, each of which had taken well-known personality tests previously. In this way, it was possible to get item analysis data and to measure the relationships of the experimental scales with other tests at the same time. Table VIII-4 shows the groups which received the tests.

Table VIII-4

Fall 1959 Experimental Administration of the Student Activities Inventory

Group	N	Grades	Sex	Supplementary Test
Youngstown, Ohio Schools	186	9-10	Both	California Psychological Inventory
Amherst College	110	14	Boys	Edwards Personal Preference Schedule
North Catholic High School, Pittsburgh	250	9-12	Boys	SRA Youth Inventory
Sacred Heart High School, Pittsburgh	190	12	Girls	SRA Youth Inventory

While it was recognized that the [illegible] were not a [illegible] sample, it was felt that

While it was recognized that the schools involved were not a random sample, it was felt that for the purposes of item analysis and estimating test interrelationships they were sufficiently like the ultimate test population to provide useful results.

Results of these experimental administrations were placed on punched cards and processed by computer. Comments of the various classroom teachers who administered the tests were collected and taken into consideration in estimating the time required, and in the preparation of directions for students and administrators for the final form.

B. Results of the Experimental Tryout

1. The Response Options

The percentage choosing each option indicated that the negative end of the 5-choice scale being used, viz., "not at all" was not being chosen. This indicated that it was too strong a term. A small study was done in which descriptive words of this type were scaled by 37 beginning college students, and the five choices best distributed were selected for the 5-choice scale in the final form. This resulted in the changes shown below:

	FROM:	TO:
A.	Very well	Very well (no change)
B.	Quite well	Well
C.	Fairly well	Fairly well (no change)
D.	Not very well	Slightly
E.	Not at all	Not very well

2. The Item Analysis

a. Procedures

For purposes of item analysis the items were dichotomized for scoring. The two extremes at the

favorable end of the scale were each scored +1, and the other three were scored -1. In the case of negatively worded items, the two extremes at the unfavorable end of the scale were scored +1, and the other three scored -1. Score on a scale was derived by taking Rights minus Wrongs. The percentage of the group choosing each option for each item, and point-biserial correlations of each item with total score on its own scale were obtained for the items on the a priori key for the 18 scales. Each a priori scale key was divided in half so that half-test scores could be correlated as estimates of scale reliabilities. Intercorrelations of all of the scales were obtained based upon the entire combined high school groups (N = 623), excluding the Amherst Sophomores. Uniquenesses* were also computed based upon the extent to which each scale score could be predicted from the multiple correlation with all of the other scales.

A preliminary analysis was done and the item analysis repeated. Scoring and item analysis were done entirely by computer.

b. Results of the item analysis

1) Based on the point-biserial correlation several items were eliminated from each scale as not being consistent with the remainder of the scale.

2) Several of the scales proved unreliable and offered little in the way of unique variances. Scale 1, Leadership, was split into Scale 1, Leadership, and

---

*Flanagan, John C. Technical Report, Flanagan Aptitude Classification Tests. Chicago: Science Research Associates, 1959. Pp. 17-22.

a new scale, 19, Bossiness, in an attempt to refine it. Similarly, Scale 8, Culture, was split into Scale 8, Manners, and Scale 20, Culture.

3) The item analysis was re-computed, producing the results shown in Table VIII-6. Refinements based upon this information gave rise to the final form. It will be noted that the reliabilities of the various preliminary scales are quite modest. However, it should be remembered that the half-test scores which were intercorrelated were based upon only 3 or 4 items in several cases.

Table VIII-6

Reliabilities and Uniquenesses

Student Activities Inventory - Fall 1959 Tryout - N = 623

Scored R - W + 200

	No. of Items	r_{11}	R^2	r_{11}*	R^2**	U^2
1. Leadership	3	.252	.167	.403	.139	.264
2. Vigor	8	.523	.326	.687	.303	.384
3. Productivity	4	.297	.392	.458	.372	.086
4. Tidiness	9	.523	.395	.687	.375	.312
5. Sociability	8	.416	.408	.588	.388	.200
6. Self-Confidence	6	.203	.095	.337	.065	.272
7. Cheerfulness	7	.321	.411	.486	.391	.095
8. Manners	8	.476	.511	.645	.495	.150
9. Talkativeness	5	.119	.166	.213	.138	.075
10. Impulsiveness	5	.197	.141	.329	.113	.216
11. Persistence	5	.227	.291	.370	.268	.102
12. Calmness	9	.556	.434	.715	.415	.300
13. Social Adjustment	7	.344	.514	.512	.498	.014
14. Theoretical	7	.179	.198	.304	.172	.132
15. Responsibility	8	.417	.517	.589	.501	.088
16. Social Sensitivity	8	.325	.317	.491	.294	.197
17. Group Centeredness	6	.300	.288	.462	.264	.198
18. Conventional	5	.066	.074	.124	.043	.081
19. Bossiness	9	.071	.086	.133	.055	.078
20. Culture	7	.390	.386	.561	.366	.195
	134***					

---

* Corrected by Spearman-Brown
** Corrected by Wherry Shrinkage Formula
*** 16 items not keyed

3. Intercorrelations of the Experimental Form with Supplementary Tests

Tables VIII-6 and VIII-7 summarize the interrelationships of the Student Activities Inventory scales with the Edwards Personal Preference Schedule, and with the California Psychological Inventory. The intercorrelations of SAI scales with the SRA Youth Inventory are essentially zero, and are thus not reported here.

a. SRA Youth Inventory

The lack of intercorrelation of SAI scales with the scales of this inventory was expected and appears reasonable. It should be noted that the Youth Inventory is simply a problem checklist, where problems are classified by surface, non-psychological type. It is reasonable, therefore, that such "scores" (number of problems in each area) would not correlate with scales developed from a much more basic psychological approach. The Youth Inventory is sometimes also scored for a 9th scale, Basic Difficulties, intended to identify students needing counseling. It appears likely that this scale (unfortunately not available for the cases in this analysis) might have correlated moderately with several of the SAI scales.

b. California Psychological Inventory

It can be seen from Table VIII-6 (see page VIII-109) that the relationships between SAI scales and the CPI scales are low, but generally in the expected direction. The low nature of these correlations appears to be due partly to the relatively low reliability of the experimental SAI scales (partially a function of their shortness), and partly to the expected differences arising from the fact that most of the

CPI scales were empirically derived from a "discrimination analysis", whereas SAI scales were derived primarily from the psychological analysis of behavior with respect to a more limited population (high school students). With these qualifications in mind, it appears that the two inventories are reasonably consistent with each other and that many of the domains being measured are similar.

c. Edwards Personal Preference Schedule

Again the interrelationships of SAI scales with the EPPS scales are low, but generally as expected (see Table VIII-7, page VIII-110). These correlations appear to be low for the same reasons as applied to those with the CPI. While the Edwards scales were developed similarly to those in the SAI, they were based upon a different psychological system (Murray's Needs). Nevertheless, as with the CPI, there is sufficient consistency and reasonableness in the relationships to support the belief that similar domains are measured.

d. Summary

The intercorrelational relationships of SAI scales with the supplementary tests were seen as lending empirical support to the construct validity of the inventory. The fact that the scales correlated at all under the conditions of the experiment was seen as a favorable factor. It is fully recognized that the SAI scales are essentially *a priori* combinations of items which

Table VIII-6

Intercorrelations of SAI Scales with California Psychological Inventory (Youngstown, Ohio, High School Students, N = 186)

	Dominance	Status	Sociability	Social Presence	Self Acceptance	Well-being	Responsibility	Socialization	Self-control	Tolerance	Good Impression	Communality	Achievmt.viaConf.	Achievmt.via Ind.	Intellect. Effic.	Psychol. Mind.	Flexibility	Femininity
Leadership	M	S	S	S	+							+						
Vigor			+					S					+					
Productivity	S	+	+				+	S		+	+		S					
Tidiness					−		+	M			−				S	$\overline{S}$	$\overline{M}$	
Sociability																−	−	
Self-Confidence	S	S	S	S	+					+	+				+			
Cheerfulness	S						+	+							+			
Manners							S	M							S	−	$\overline{S}$	
Talkativeness				+	+			−	$\overline{S}$									
Impulsiveness	+			+	+				−									
Persistence					$\overline{S}$	S		+							S	$\overline{S}$	$\overline{M}$	
Calmness	+	+				M	M	M	S			+	S		S			
Social Adjustment			+			S	S	M					S		M		$\overline{M}$	
Theoreticality																		
Responsibility							+	M					S				$\overline{S}$	
Social Sensitivity							S	S							S	−	$\overline{S}$	
Group-Centeredness							+	S										
Conventionality	−				−					$\overline{S}$			−	−		−	$\overline{S}$	
Bossiness						$\overline{S}$			$\overline{S}$									
Culture							+											

Note: + or − indicates r .15-.18 in the plus or minus direction respectively (significant at the two-tailed 5% level)

S or $\overline{S}$ indicates r of .19-.24 in the plus or minus direction respectively (significant at the two-tailed 1% level)

M or $\overline{M}$ indicates r of .25-.40 in the plus or minus direction respectively.

Table VIII-7

Intercorrelations of SAI Scales with
Edwards Personal Preference Schedule
(110 Amherst College Sophomores)

	Achievement	Deference	Order	Exhibition	Autonomy	Affiliation	Intraception	Succorance	Dominance	Abasement	Nuturance	Change	Endurance	Heterosex.	Agression	Consistence Score
Leadership		-							+							+
Vigor																
Productivity			+			-							M			
Tidiness			M		$\bar{S}$											
Sociability	-				-	+										
Self-Confidence																
Cheerfulness						+				-			S		-	
Manners																
Talkativeness		-		+												
Impulsiveness																
Persistence			M					-					M			
Calmness																
Social Adjustment	-				$\bar{S}$								+			
Theoreticality			-				S		-							
Responsibility			S										M			
Social Sensitivity					-						M				-	
Group-Centeredness	-												+			
Conventionality					$\bar{M}$	+	-									
Bossiness														-		
Culture												+	-			

Note: **+ or - indicates r of .19-.24 in the plus or minus direction, respectively (significant at the two-tailed 5% level)**

S or $\bar{S}$ indicates r of .25-.29 in the plus or minus direction, respectively (significant at the two-tailed 1% level)

M or $\bar{M}$ indicates r of .30-.45 in the plus or minus direction, respectively.

do not all measure the same thing to the same degree. However, it was demonstrated that the scales were measuring the desired domains of personality, and that the refined and consolidated final form could be expected to yield more reliable scale correlations.

C. The Final Form

On the basis of the item analysis, intercorrelational data, and the uniquenessess and reliabilities of the various scales, several scales were omitted, others were combined, and about half of the items dropped. Items which showed poor item response distribution, high percentage of "omits", appeared ambiguous or otherwise undesirable, or failed to relate to one of the original 20 scales, were dropped. In this way the number of items was reduced to 150 for the final form.

While 150 items were retained in the final form, only 13 scale scores could be obtained, due to practical limitations in scoring and analysis. The 13 best of the possible scales were selected for scoring. For the "Final Form", scales with low uniqueness but relatively high and consistently similar patterns of intercorrelations were combined, so that the additional items would provide increased scale reliability. Those with low reliability, low uniqueness, and no such pattern, were dropped. In one or two cases, scales were retained on the basis of their psychological potential. The 13 scales scored in the final form are show, along with the combinations which were made, in Table VIII-8 (page VIII-112).

TABLE VIII-8

Scales Scored, SAI Final Form

Scale	Number of Items	Experimental Scale from Which Derived
1. Sociability	12	Sociability and Cheerfulness
2. Social Sensitivity	9	Social Sensitivity
3. Impulsiveness	9	Impulsiveness
4. Vigor	7	Vigor
5. Calmness	9	Calmness
6. Tidiness	11	Tidiness
7. Culture	10	Culture and Manners
8. Leadership	5	Leadership
9. Self-Confidence	12	Self-Confidence
10. Mature Personality	24	Productivity, Persistence and Responsibility
11. Conventionality	7	Conventionality
12. Theoreticality	8	Theoreticality
13. Group-Centeredness	6	Group-Centeredness
Unscored	21	All scales

It is not expected that the 13 scales scored in the final form cover the range of high school personality. They represent, rather, the best-measured of the initial group of scales. However, effort will be made to retain the much broader coverage of the 150 individual items themselves, if feasible, to relate to student behavior. Thus it is planned not to rely solely on the scale scores in this area.

Appendix 8 b

The Description and Development of the Interest Inventory

General Purpose:

The general purpose of the Interest Inventory was to obtain direct expressions of interest (like versus dislike) toward occupations representing a broad range of fields, differential ability requirements, differential educational requirements, etc. Similarly, expressions of interest were desired on activities including vocational activities, sports, school subjects, hobbies, pastimes, etc.

It was the feeling of the advisory groups and the staff that an inventory which approximately paralleled accepted instruments in present use should be included in order that further empirical tests could be made of their predictive power and their usefulness in guidance. The objective, the Panel felt, was to gather a varied set of items into the study so that these could be combined later in whatever ways would be desirable to reproduce fairly closely the dimensions of several of the existing inventories.

Description of the Inventory:

A. Definition of an interest. The commonly used interest inventories define "interest" in terms of expressed liking or preference for various occupations, activities, school subjects, types of people, social situations, etc. This is the definition of "interest" intended here. Other definitions of "interest" have been made in terms of amount of information retained in given subject matter areas (see

the Information Test) and in terms of overt activities (see the Student Information Blank).

Inventoried interests are subjective as contrasted with the objectivity of interests measured by information and by reports of overt activity. Furthermore, inventoried interests are not independently verifiable or observable. Inventoried interests may be said to be molar, subjective, dispositions of the respondent toward common activities and experiences. The specific activities and experiences included in this inventory are those familiar to a high school population and relevant to the project; i. e., occupations, school subjects, and activities.

B. Inventory directions. The purpose of the directions to the inventory was to describe, define, and clarify the task of the respondent; the context in which the items are to be considered; and to motivate the individual to respond in a manner which truly reflects his feelings. In this task we followed closely upon precedents established by Strong and Kuder.

The respondent's task is to indicate the degree to which he likes, dislikes, or is indifferent to a variety of occupations and activities. In order to clarify the context in which the items are to be considered, the respondent is told to consider only the activity involved in the occupations; not to consider associated factors such as salary, social standing, permanence of the work, training or education needed, etc. Particular emphasis was placed upon responding in terms of the degree of his liking or disliking the activity involved. The directions read as follows:

"This section contains items for you to answer about

occupations you would like and things you would like to do. For each item assume that you would have any necessary training or education that would be required. Disregard salary, social standing, permanence, etc., in fact anything except how well you would like to do the work or the activity. Your answer does not mean that you plan to go into an occupation if you say you would like it--only that it involves the kind of activity you think you would enjoy."

C. Item format. Item stems were presented in the briefest, simplest form possible. It was the feeling of the staff that statements, phrases, or complete sentences are unnecessary if the same content can be contained in a single word. Further, brief items would be less likely to pose a reading difficulty problem, would be likely to consume less reading time, and would take up less space in the test booklets.

For each item, the respondent was asked to indicate the degree to which he would like or dislike the occupation or activity on the following scale:

a. I would like this very much.
b. I would like this fairly well.
c. Indifferent or don't know much about it.
d. I would dislike this a little.
e. I would dislike this very much.

Developing the Item Pool:

The items in the Strong Vocational Interest Blank, the Kuder Preference Record, the Thurstone Interest Schedule and the Minnesota Vocational Interest Inventory were typed on cards and the content analyzed logically to reduce the redundancy among the items. The content analysis involved

sorting items by occupational titles and by activities. Activities whose content could be sorted with occupational titles were so sorted. Thus the final list consisted largely of occupational titles and occupational activities, and a short list of activities not clearly associated with occupations.

Several preliminary versions of the Interest Inventory were presented to the advisory panels for review and comment. The recommendations of the Panel members and the judgment of the staff were used to reduce the number of items to a manageable number.

As a further check on the adequacy of coverage, the reduced set of occupational titles was sorted with the Bureau of the Census occupational classification. The purpose of this check was to assure adequate (though not necessarily proportional) representation of each Bureau of the Census category. Another check was made to determine whether representation in the Interest Inventory was adequate for those occupations in each Bureau of the Census Classification that are most frequent in the population (1950 figures). A small number of items, both occupations and activities, were added to the inventory as a result of this check.

The Pilot Study

In the fall of 1959 an experimental form of the Interest Inventory was prepared for use in a pilot study. The experimental form contained 300 items, of which 164 were occupational titles and 136 were activities.

The objectives to be served by the pilot study were several in number. Homogeneous scales were to be developed and correlated with existing inventories to determine the

degree to which the Interest Inventory accounted for the reliable variance in these inventories. An estimate was needed of the amount of time the students would require to complete the form. The Inventory was to be untimed in the sense that everyone would be allowed to finish. This estimate was needed in order to apportion time allowances in the battery among the tests and inventories. The pilot study was also expected to uncover any unanticipated difficulties in the administration or with the Inventory itself.

A. Sample and Administration. The sample included two high schools in Texas and a group of entering Freshmen at the University of Pittsburgh. The experimental form of the Interest Inventory was administered to the entire sample of 655 students. Some students had taken the KPR and others were given the SVIB. Table ~~1~~ VIII-9 presents the breakdown of the number of students in each school and grade to whom the Kuder Preference Record and the Interest Inventory were administered, the number to whom the Strong Vocational Interest Blank and the Interest Inventory were administered, and the total number of cases in each grade.

TABLE VIII-9

Number of Cases with the Interest Inventory, Kuder Preference Record and Strong Vocational Interest Blank

Inventories Administered	School and Grade	N
Interest Inventory	Univ. Pgh. Freshmen	270
	School A - 12	124
	School A - 10	163
	School B - 9	98
	Total N -	655
Strong V.I.B.	Univ. Pgh. Freshmen and School A-12	415
Kuder P.R.	School A - 12, A - 10	163

Administration of the Interest Inventory was conducted under the guidance of persons experienced in testing procedures. Administrators were provided with a set of directions on which they were asked to indicate the amount of time taken by their group to complete the Inventory and to comment upon any unusual circumstances or problems in testing. At five minute intervals the students were instructed to circle the number of the item on which they were working.

B. Construction of the Homogeneous Scales. A priori scales were constructed based upon the specifications for the Inventory, a review of several factor analytic studies, and the judgment of the staff. Correlations between the items and the scale total score was used for the empirical work on the scales. From the items allocated to each scale a smaller number were designated as "core" items. "Core" items were those that previous research had demonstrated most clearly to be homogeneous. For example, "core"

items in the Science Scale included occupations such as astronomer, biologist and chemist. A total score for each scale was obtained from the "core" items alone. The scoring formula was R-W+200, in which options A and B were scored R, and C, D, & E, were scored W. The addition of the constant avoided the analytic complications of negative scores. Each item was correlated with the scale total score. The scales were purified by rejecting those items that did not correlate significantly with the total score.

C. Results of the Pilot Study. The data was analyzed on the IBM 704 Electronic Computer which scores and analyzes simultaneously. Item analysis statistics included the percent selecting each item option, the percent omitting the item, the scale score mean for those selecting each item option, the point biserial correlation coefficient between each item option and total score on the scale, and the mean and standard deviation for each scale. This data was used to judge the value of items for the Inventory and for the scales in which they were placed. The Inventory was reduced from 300 to 205 items on the basis of this data.

In addition to the item analysis data, correlation matrices were obtained for the Interest Inventory scales separately, and with the Kuder PR and Strong VIB. The correlation matrices are presented in Tables VIII-10, VIII-11, and VIII-12.

Table VIII-10

Interest Inventory Scale Intercorrelations (N=655)

	Scales	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1.	Mech. Tec.	85	52	51	17	28	30	-06	14	43	24	42	-09	-01	-04	46	05
2.	Skilled Trdes	52	71	10	08	28	20	36	32	25	26	42	14	13	21	69	28
3.	Science	51	10	93	34	22	25	-26	-01	32	25	26	10	11	08	05	16
4.	Public Serv.	17	08	34	84	46	29	-05	31	13	32	16	33	22	19	11	32
5.	Business Mgt.	28	28	22	46	56	35	15	49	18	35	20	28	16	18	32	35
6.	Computation	30	20	25	29	35	65	27	29	06	18	09	12	02	03	23	23
7.	Office Work	-06	36	-26	-05	15	27	83	35	-14	07	-06	25	13	26	28	32
8.	Sales	14	32	-01	31	49	29	35	74	07	23	14	28	20	24	34	40
9.	Outdoor Rec.	43	25	32	13	18	06	-14	07	67	25	48	-12	-09	-10	29	-02
10.	Sports	24	26	25	32	35	18	07	23	25	48	19	18	14	12	26	23
11.	Farm, Ranch	42	42	26	16	20	09	-06	14	48	19	75	02	05	04	36	14
12.	Lit.-Ling.	-09	14	10	33	28	12	25	28	-12	18	02	78	55	63	03	45
13.	Musical	-01	13	11	22	16	02	13	20	-09	14	05	55	81	55	00	35
14.	Artistic	-04	21	08	19	18	03	26	24	-10	12	04	63	55	72	00	39
15.	Labor	46	69	05	11	32	23	28	34	29	26	36	03	00	00	77	21
16.	Social Serv.	05	28	16	32	35	23	32	40	-02	23	14	45	35	39	21	55

Note: Decimal points omitted. Principal diagonal contains the split-test reliability corrected by the Spearman-Brown formula.

Table VIII-13 presents half-test reliability coefficients, squared multiple correlation coefficients, and uniqueness coefficients for the Interest Inventory based on 655 cases. The reliability coefficients appear to indicate that the level of reliability is acceptable for most of the scales.

The uniqueness coefficients indicate that each scale is contributing variance that differs from the other scales. Tables VIII-14 and VIII-15 present the multiple correlations and shrunken multiple correlations of each Interest Inventory scales with each of the Kuder PR scales and eleven selected Strong VIB scales (the Strong scales were selected before analyzing). The size of these coefficients is adequate in most cases.

Table VIII-11

Interest Inventory
and Kuder Preference Record Intercorrelations
(N=163)

II Scales	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Mech. Tech.		42	49	18	33	23	-24	21	53	35	44	-14	11	-07	50	02
2. Skilled Trdes	42		02	08	37	04	22	36	26	27	41	17	24	29	70	26
3. Science	49	02		30	32	32	-32	05	44	36	24	02	01	00	08	13
4. Public Serv.	18	08	30		39	28	-05	23	31	29	29	26	17	16	12	14
5. Business Mgt.	33	37	32	39		34	18	46	28	39	23	28	18	25	39	29
6. Computation	23	04	32	28	34		25	32	09	15	10	27	15	10	17	23
7. Office Work	-24	22	-32	-05	18	25		43	-24	01	-11	35	14	28	17	38
8. Sales	21	36	05	23	46	32	43		16	22	11	30	18	22	30	33
9. Outdoor Rec.	53	26	44	31	28	09	-24	16		38	52	-12	-04	-09	42	-01
10. Sports	35	27	36	29	39	15	01	22	38		35	14	08	08	33	12
11. Farm, Ranch	44	41	24	29	23	10	-11	11	52	35		-04	00	10	39	04
12. Lit. -Ling.	-14	17	02	26	28	27	35	30	-12	14	-04		45	59	03	44
13. Musical	11	24	01	17	18	15	14	18	-04	08	00	45		38	17	29
14. Artistic	-07	29	00	16	25	10	28	22	-09	08	10	59	38		05	35
15. Labor	50	70	08	12	39	17	17	30	42	33	39	03	17	05		14
16. Social Serv.	02	26	13	14	29	23	38	33	-01	12	04	44	29	35	14	
KPR Scales																
17. Outdoor	40	19	38	11	-02	-04	-37	-17	60	19	60	-23	-12	-13	17	-13
18. Mechanical	70	27	39	15	12	-01	-41	-04	54	27	40	-40	-12	-30	37	-26
19. Computation	-01	-28	24	17	10	48	-04	00	-08	03	-15	-02	-11	-24	-16	-02
20. Scientific	36	-08	67	26	04	13	-45	-14	40	21	21	-21	-17	-27	-03	-09
21. Persuasive	01	04	-18	05	18	-12	14	36	-11	-02	-25	-06	-03	-05	06	-01
22. Artistic	-09	12	-07	-07	-09	-14	-03	-05	-07	02	-01	08	06	37	-06	-06
23. Literary	-20	00	-13	06	09	06	20	20	-16	-05	-22	45	06	20	-06	23
24. Musical	-12	09	-29	-07	08	-01	23	10	-20	-09	-16	28	62	30	02	13
25. Social Serv.	-36	-08	-21	-18	-19	-06	31	-05	-24	-11	-10	13	-06	09	-11	29
26. Clerical	-45	-21	-45	-10	-02	16	52	12	-44	-22	-36	16	00	07	-19	16

Note: Decimal points omitted

Table VIII-12

Interest Inventory and Strong Vocational Interest Blank Intercorrelations (N=415)

II Scales	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Mech. Tech.		55	51	18	27	36	10	14	40	23	42	-06	-04	-02	46	07
2. Skilled Trades	55		15	09	27	24	45	34	26	26	43	15	12	20	69	31
3. Science	51	15		36	17	25	-17	-05	28	23	27	16	17	11	04	16
4. Public Serv.	18	09	36		44	29	03	29	09	36	10	34	22	18	11	36
5. Business Mgt.	27	27	17	44		34	21	49	17	36	18	26	13	15	32	34
6. Composition	36	24	25	29	34		28	27	07	19	06	05	-03	-03	25	20
7. Office Work	10	45	-17	03	21	28		40	-04	10	00	25	14	26	35	36
8. Sales	14	34	-05	29	49	27	40		06	25	13	25	18	21	37	43
9. Outdoor Rec.	40	26	28	09	17	07	-04	06		21	48	-08	-08	-08	28	00
10. Sports	23	26	23	36	36	19	10	25	21		16	18	13	11	26	26
11. Farm and Ranch	42	43	27	10	18	06	00	13	48	16		01	04	-02	36	15
12. Lit. -Ling.	-06	15	16	34	26	05	25	25	-08	18	01		56	64	03	44
13. Musical	-04	12	17	22	13	-03	14	18	-08	13	04	56		58	-04	34
14. Artistic	-02	20	11	18	15	-03	26	21	-08	11	-02	64	58		-02	37
15. Labor	46	69	04	11	32	25	35	37	28	26	36	03	-04	-02		24
16. Social Serv.	07	31	16	36	34	20	36	43	00	26	15	44	34	37	24	
SVIB Scales																
17. Doctor	16	-15	51	29	00	-01	-38	-20	06	06	05	15	21	15	-29	-03
18. Engineer	39	-10	53	15	-02	13	-39	-30	16	-03	06	-17	-05	-12	-19	-23
19. Prod. Manager	32	-10	41	34	24	21	-30	-07	16	09	06	-06	-02	-10	-11	-10
20. Farmer	31	-11	40	22	09	12	-40	-15	17	06	14	-14	-02	-12	-14	-11
21. Carpenter	45	07	36	08	04	20	-22	-15	20	08	17	-18	-09	-06	-03	-11
22. Office Worker	02	-11	11	38	30	30	-04	23	-09	15	-02	18	12	14	-14	16
23. Pub. Adminis.	-12	-26	13	04	-14	-17	-38	-24	-11	-11	-10	12	18	21	-37	-12
24. Soc. Sci. Teacher	09	-13	30	49	36	13	-23	08	04	17	09	24	21	12	-15	18
25. Lawyer	-07	-15	09	43	34	08	-13	24	-10	16	-02	32	24	20	-15	30
26. Life Ins. Sales.	-24	-38	05	34	15	-07	-41	01	-21	03	-19	21	17	09	-34	02
27. Musician	17	-30	-03	30	29	03	-33	18	-14	05	-16	15	11	06	-24	01

Note: Decimal points omitted

TABLE VIII-13

Interest Inventory

Reliability and Uniqueness Coefficients
(N=655)

		r_{11}	r_{11}*	R^2	R^{2**}	U^2
1.	Mechanical-Technical	.737	.849	.584	.574	.275
2.	Skilled Trades	.553	.712	.635	.626	.086
3.	Science	.873	.932	.479	.467	.465
4.	Public Service	.727	.842	.382	.367	.475
5.	Business Management	.389	.560	.424	.410	.150
6.	Computational	.483	.651	.299	.283	.368
7.	Office Work	.704	.826	.432	.419	.407
8.	Sales	.582	.736	.394	.380	.356
9.	Outdoor Recreation	.500	.667	.357	.342	.325
10.	Sports	.315	.479	.227	.208	.271
11.	Farm and Ranch	.601	.751	.367	.352	.399
12.	Literary-Linguistic	.643	.783	.536	.525	.258
13.	Musical	.680	.810	.388	.374	.436
14.	Artistic	.566	.723	.495	.483	.240
15.	Labor	.628	.771	.556	.545	.226
16.	Social Service	.379	.550	.385	.370	.180

* Corrected by Spearman-Brown formula.
** Corrected by Wherry's Shrinkage formula.

TABLE VIII-14

Multiple Correlation Coefficients

for the Interest Inventory with the Kuder PR

		R	R*
1.	Outdoor	.79	.76
2.	Mechanical	.81	.78
3.	Computational	.67	.63
4.	Scientific	.78	.75
5.	Persuasive	.63	.59
6.	Artistic	.52	.45
7.	Literary	.70	.66
8.	Musical	.64	.59
9.	Social Service	.54	.47
10.	Clerical	.73	.70

* Corrected by Wherry's Shrinkage Formula

TABLE VIII-15

Multiple Correlation Coefficients
for the Interest Inventory with Selected Scales
from the Strong VIB

	Strong VIB	R	R*
1.	Doctor	.66	.64
2.	Engineer	.75	.74
3.	Production Manager	.64	.62
4.	Farmer	.58	.55
5.	Carpenter	.59	.57
6.	Office Worker	.58	.56
7.	Public Administrator	.56	.53
8.	Social Science Teacher	.72	.71
9.	Lawyer	.62	.60
10.	Life Insurance Salesman	.73	.71
11.	Musician	.88	.87

* Corrected by Wherry's Shrinkage Formula

CHAPTER IX

THE GENERAL SCHOOL CHARACTERISTICS QUESTIONNAIRE

Rationale

The preceding chapters have outlined some of the vast amount of information about each student that has been collected by Project Talent. However, children are a product of many influences, not the least of which are the schools they attend. Thus, in order to understand more fully the background and abilities of the children themselves, it was necessary also to look at their schools directly. The present chapter deals with one of the instruments which was developed for this purpose, the General School Characteristics Questionnaire.

As its name implies, this questionnaire was intended to provide information about a variety of characteristics of the school which might make a difference in its students as compared to those of other schools. It was filled out in each school by a person well qualified to provide information about the school as a whole, usually the school principal.

One of the most important usages of such school information will be as a control factor in interpreting student data. For example, suppose we wish to study the achievement test scores for seniors who plan to be scientists. Many school characteristics could affect these scores. Obviously, the number and kinds of science and math courses offered by the school would affect the achievement scores made by its pupils. Many other factors could be important and have to be examined before a valid conclusion can be drawn. These might include: size of school; size of instructional classes; percentage of male teachers; size of community; location of community; characteristics of the area served by the school; number of grades in the school; amount of graduate training of teachers; years of teaching experience of staff and administrative experience of principal; guidance facilities available; extra-curricular activities available; grouping, tracking, and acceleration-promotion policies, etc.

It may be that only a few, or even none, of these factors has any measurable bearing on the level of achievement test scores shown by seniors planning to be scientists. The important thing, however, is that the availability of such information about the schools makes it possible to study and measure the relationship of these factors to the abilities, achievements, and behaviors of students.

Of course the example given above is only one of many uses to which school data can be put. We are interested in studying the long-range effects of various school policies and characteristics on career choices, satisfactions, and successes; in studying the types of schools that develop in different types of communities; and in knowing some of the

characteristics of the nation's teachers.

We are also interested in the characteristics of our schools themselves: How old is the average school building? How large is the average class in city schools, country schools, in the North, in the South? What percentage of the high schools offer courses in Russian? How many schools have science labs? Etc.? Through its large and representative sample of the nation's high schools, Project Talent has worked toward providing the best and most complete information on these and other topics that has ever been available.

Choice of Topics

From the very beginning, the General School Characteristics Questionnaire seemed destined to be of monumental length. As the staff began to review the literature and to develop the actual questions, the list grew and grew--and grew some more. When a preliminary form the lengthy questionnaire was submitted to the Educational Research Panel, charged with the responsibility for defining appropriate educational issues for study, it responded by producing an even more lengthy list of recommended areas for study. These areas are summarized below, under their major categories:

A. Characteristics of the school which may influence educational achievement and decisions.

1. Culture of the school and its community.
2. Physical features of the school and its community.
3. School policies and practices.
4. Content and procedures of courses.

B. Characteristics of the students which may influence educational achievements and decisions (in addition to those characteristics measured by the basic test battery).

1. Physical characteristics.
2. Culture of his home.
3. Choice of alternatives in the school program open to him.
4. Motives, values, drives, purposes.
5. Educational, social, civic, and employment experiences.

Subsequent study and refinement of the list of suggested areas resulted in the decision to place many of these areas in different instruments. Areas concerning the attitudes and personal background of the student became parts of the Student Information Blank, discussed in the preceding chapter; separate questionnaires (discussed in the next chapter) were developed to cover the guidance program of the school, and the characteristics of its counselors.

At each of the later meetings of the Educational Research Panel, additional areas for inclusion in the General School Characteristics Questionnaire were proposed; methods were discussed and topics for elimination were suggested. It should be noted, however, that the proposals for addition continued to out-number those for elimination. For example, the following areas were proposed for inclusion at the September 1958 meeting:

1. Fiscal characteristics of the school, including various indices relating such factors as college attendance, ability to support education, and socio-economic or cultural status to educational expenditures.

2. Teacher-pupil ratio.

3. Rural-urban classifications, possibly along census lines.

4. Some measures of student leadership within the school, the extent of student participation in planning educational experiences, the activity program, student government, and the like.

5. Measures of the holding power of the school and of the age ranges and average ages for each grade.

6. Some estimates of community attitudes possibly taken indirectly from pupil attitudes.

7. Some measures of educational change in the school and community.

8. Identification of certain kinds of talent, including the artist, musician, writer, athlete, school leader, etc.

Due to the increase in number of areas to be considered, one of the major problems in the development of the questionnaire was pruning the number of questions asked to a number which a principal might reasonably be expected to answer. It was decided that questionnaire form would provide the greatest versatility, and that most questions should be placed in multiple choice form in order to minimize the principal's task. After some pruning, the staff produced an extensive questionnaire for the review of the Educational Research Panel. Recommendations for further reduction in length were requested, and many general recommendations and specific suggestions on all aspects of the questionnaire were received from the Panel.

For example, it was suggested that more questions be added on community facilities, the questions on school

philosophy be revised or dropped, and the section on courses offered in the school be revised. Additional fiscal data were recommended--per-pupil expenditure, library facilities and expenditures, etc.

The Development and Tryout of the Experimental Form

Based upon staff research and the reviews and suggestions of the Educational Research Panel, a tryout questionnaire was prepared. This questionnaire was sent to seven high school principals in several Georgia communities for tryout. While this was not intended to be a representative sample, it was set up to cover large and small schools of both rural and urban character. The comments, criticisms, and suggestions of these principals and of the individual Panel members, along with further study by the central staff, formed the basis for a new revision of the questionnaire.

The newly revised questionnaire was submitted to the Panel for a final review at its meeting in October 1959. At this time the Panel was asked to submit detailed suggestions with respect to form of the questionnaire. The Panel approved with some suggestions the general content and form of the questionnaire, including the proposed intervals for the various item choices. The recommendations included proposals for revising the order of the questions, revising the course listings, and some specific suggestions regarding particular items.

The suggestions and recommendations of the October Panel meeting were then incorporated with the revised questionnaire to make the final form of the questionnaire.

response such as number of school days in the school year, per-pupil expenditure, and enrollment data.

In the course listing section, courses were grouped by broad subject matter areas. For each of the courses listed, principals were asked to write one of the following code numbers in the parentheses in front of the course:

4 -- This course is offered and is compulsory for all students.

3 -- This course is offered and is compulsory for for all college preparatory students, but not for all students.

2 -- This course is offered and is compulsory only for some group of students other than college preparatory students.

1 -- This course is offered as an elective only.

0 -- This course is not offered.

A Picture of Three Schools. The final Questionnaire is too long to reproduce here, but the kinds of information included can be illustrated by comparing some of the actual data supplied by three different schools. The three examples given below are not presented as being typical or representative, but were simply chosen essentially at random from the Project files for the sake of illustration. These examples show the type of descriptive picture which the Questionnaire material provides. These descriptions do not include all of the information available about these three schools but illustrate the areas of coverage contained in the Questionnaire.

School 1. This is a large, metropolitan public school with a total enrollment for grades 9, 10, 11, and 12 of 1795 boys and 1880 girls. It is a fully accredited four-year high school operating on a school year of 192 days, each composed of seven 45-minute periods. Classes average

The Final Form

The following section presents a brief description of the final form of the General School Characteristics Questionnaire.

Content. In general, the final form of the General School Characteristics Questionnaire is arranged in a number of sections. The first section deals with the school and its policies, practices, and plant. Questions included here concern type of school, special classes, grading and advancement policies, condition of the plant, average class size, etc. The next section deals with number, training, and characteristics of the teaching staff. Then come questions pertaining to the students: enrollment, dropout, college-going, remedial work, etc. Next is a section concerning characteristics of the community, such as PTA activity, type of area served by the school, per-pupil expenditures, ethnic background, and community facilities. The final section of the main Questionnaire concerns the age, experience, and training of the principal himself. At the end of the Questionnaire is an extensive listing of high school courses where the principal is to indicate the courses offered in his school.

Format. Virtually all of the Questionnaire is placed in the multiple choice form so that the principal can indicate his answer to the question by marking one (or in some cases several) of the indicated choices. A number of questions provide an opportunity to write in and explain an "Other" response not covered by the indicated choices.

A very few questions were set up as fill-in questions. These are confined to questions requiring a numerical

33-38 students, who are assigned 2-3 hours of homework per day. This school provides several kinds of recognition for superior work; an advanced curriculum in all courses for superior students; and special separate classes for mentally retarded, reading and speech difficulties, rapid learners, and special talents. Numerical grades are assigned primarily on level of subject matter achievement and are used as the principal basis for assigning students to instructional classes. Summer school and adult education courses are offered. Enrollment was down somewhat over the preceding year, daily absenteeism averages 3-5%, and the Administrative Assistant who filled out the Questionnaire for review by the principal felt that the school needs additional personnel in administrative, supervisory, guidance, teaching, and clerical positions.

The school building is 20-24 years old but was renovated less than three years ago. It is equipped with a library of more than 2700 books, cafeteria, gym, athletic field, auditorium, movie, radio, television, phonograph and public address system equipment, various laboratories, etc. A wide variety of extra-curricular activities and clubs are available.

All of the more than 27 teachers are full-time, and 50-59% of them are men. Virtually all of the teachers have a Master's degree, all are fully state-certified, and are spending at least half-time teaching in their major areas of preparation. Beginning teachers start at $4500-$4900 per year, turn-over last year was 3-5%, and the current staff averages more than 15 years of full-time teaching experience.

Transfers last year account for less than 4% of the

enrollment; less than 9% of the boys and girls entering 10th grade drop out before graduation, while 70-79% go on to college. Delinquency rate is about 1%, and less than 4% of the students take remedial math or remedial reading or English, and only 5-9% take repeat work in summer school.

While no citizens' group has studied the school recently, more than 63% of the parents are members of the PTA, which meets about once every five weeks. The area served by the school is about equally apartments and homes in a moderately expensive urban residential area. Per-pupil expenditures last year were about $560. Many community facilities such as concerts, theater, library, museum, Scouts, swimming, etc., are readily available.

The principal of this school is a man of 55-59, with 15-19 years experience as a principal; he has been principal of this school 6-8 years. He holds a Doctor's degree and has had more than 50 credits in education courses.

The courses offered in this school cover a wide range of commercial and college preparatory subjects; a few vocational courses and shops are also offered.

School 2. This is a small, midwestern public school with a total enrollment for grades 9, 10, 11, and 12 of 96 boys and 107 girls. It is a state-accredited 12-year school, operating on a school year of 178 days composed of 8 or more 45-minute periods each day. Classes average 24-29 students, who are assigned 1-2 hours of homework a day, and have 9 or more study halls per week. This school provides special-group, separate classes only for blind students, has no accelerated curricula, and recognizes superior achievement only through a "Dean's List" or equivalent, and the

privilege of taking additional work. Numerical grades are assigned primarily on the basis of subject matter achievement and are used in combination with tests and staff judgment to make class assignments. There is no summer school program, and only non-credit vocational courses are offered for adults. Enrollment is about the same as last year, with 3-5% average daily absenteeism. Personnel needs include supervisory, guidance, clerical, and janitorial.

The school building is 5-9 years old and received painting or renovation less than 3 years ago. It has a library of 1500-1799 books, cafeteria, gym, athletic field, and auditorium, movie, radio, phonograph and public address equipment, science labs, etc. Extra-curricular activities are not as broad and varied as for School 1 but include intramural athletics (boys and girls), band, school papers, dramatics, subject matter clubs, dances, and boys' interschool athletics.

In contrast to School 1, there is no homogeneous grouping and no acceleration is permitted; however, like School 1, there are several tracks. There are 9-11 full-time teachers, of whom 70-79% are men. All of these teachers have a Bachelor's degree and none has any higher degree, only one or two having any graduate training. All of the teachers are state-certified, and average 15 years or more of full-time teaching experience. Beginning teachers start at $3500-$3999 per year, and turn-over last year was zero.

Transfers last year were less than 4%, while drop-out from 10th-12th grade averaged 10-19%. Only 10-19% of the boys and 0-9% of the girls go on to college; delinquency

rate was about 1%.

No citizens' group has recently studied the school, and only 7-13% of the parents are members of the PTA, which meets monthly. The school serves a rural farm area in which the residences are primarily moderate-priced homes about 25-29 years old. Per-pupil expenditures last year were about $233. Community facilities include library, 4-H, Scouts, and public recreation center.

The principal of the school is a man of 50-54 who has had all of his teaching and administrative experience (15-19 years) in this school. He holds a 4-year degree with 50 or more credits in education.

The courses offered are those of a comprehensive high school but are considerably less varied than those of School 1. No fine arts are offered and Latin is the only language offered, as compared to Spanish, French, Latin, and Hebrew in School 1. Four agriculture courses are offered, however.

<u>School 3</u>. This is a moderate sized, <u>non-public</u> school located in the largest city of a western state. It has a total enrollment of 196 boys and 352 girls in the top four of its 12 grades. It is fully accredited and operates on a 177-day school year, with 6 periods of 55 or more minutes a day. Classes average 27-29 students, who are given 2-3 hours per day of homework and no study halls. Special separate classes are provided only for reading difficulty and math difficulty; and "Dean's List", special awards, and accelerated math, science, and language curricula are provided for superior students. Numerical grades are assigned on primarily the basis of subject matter achievement

and are used with test and staff judgment to make class assignments. Students can take both new and repeat work in summer school, but no adult education is offered. In contrast to Schools 1 and 2, religious education is contained in the regular curriculum. Enrollment is about the same as last year, with daily absenteeism averaging 6-8%. Administrative, supervisory, guidance, and clerical personnel are needed.

The building is less than five years old and is equipped with a library of more than 2700 books. It has essentially the same equipment as Schools 1 and 2. Extra-curriculars include school government, magazine, athletics, subject matter clubs, glee club and band, drama, debate, and religious club.

Many courses are grouped homogeneously and there are several tracks, but no acceleration is permitted.

There are about 21-23 full-time teachers and 9-11 regular part-time teachers; only 10-19% of the full-time teachers are men, and about 12-14 of the full-time teachers have a Bachelor's degree, while 9-11 have a Master's degree. About 90-99% are state-certified, and only 50-59% are teaching in their major area of preparation. Starting salaries are $3500-$3999 per year, and average years of full-time teaching experience are 12-14.

Transfers last year accounted for 5-9% of enrollment; both delinquency and drop-out rates were nil. However, only 20-29% of the boys and 40-49% of the girls go on to college. Less than 4% of the students take repeat or remedial summer work, but 5-9% are in remedial math and 10-14% in remedial reading or English.

No citizens' group has recently studied the school, but more than 63% of the parents are members of the monthly PTA. Pupils attending the school come from areas scattered over the entire city and from all types of homes of all ages. Per-pupil expenditure last year (not comparable to public schools for various reasons) was reported as $100. Public facilities include library, museum, concerts, and Scouts.

The principal of this school is a woman 35-39 years of age who has had 15-19 years of teaching experience, less than 5 years as principal, and is in her first year as principal of this school. She holds a Master's degree and has had 40-49 credit hours of education courses.

Courses offered are those of a general comprehensive high school. Latin, French, and Spanish are given, but very little is offered in industrial arts and nothing in trade and vocational education. Several religion courses are offered, as are music and fine arts.

A Final Word. It was necessary to develop a single questionnaire to be used for all schools in the study. In practice, some problems resulted from this approach. For example, some questions in the final form were well phrased for junior high schools; in some cases, schools in some of the special testing were not secondary schools at all. However, a great majority of these difficulties have been eliminated by contacting a few schools for more information and by careful examination and coding of responses. When necessary, the aid of the Regional Coordinators was used to help complete the questionnaires and clarify them after they were initially filled out.

Finally, it should be noted that, broad as the

information requested in the Questionnaire may be, it forms only a small part of those important questions which might have been asked. As pointed out before, the most difficult task which faced the staff (and the Panel) concerning this Questionnaire was not what to include, but what to exclude. Decisions were made in this respect on the basis of those items of information which appeared most relevant to the analysis of the other data being collected; those which seemed most important in their own right; and those which appeared to be reasonably accessible to the school principal. It is unfortunate but unavoidable that the data for many important studies were thus left out in order to make room for data considered more important or more appropriate by the Panel and the staff.

CHAPTER X

THE GUIDANCE PROGRAM QUESTIONNAIRE AND THE COUNSELOR'S QUESTIONNAIRE

Early in the planning stages of the project a single questionnaire was considered which included questions on guidance and counseling as a part of the general questionnaire to be completed by the principal at each school. This approach was soon abandoned upon the recommendations of our advisory panels (Educational Research Panel and the Guidance and Counseling Panel). There were several reasons for this. First, the volume of information that was wanted from each school seemed to be too large to be answered by the principal. Therefore, the staff and panels felt that many questions might be retained if someone other than the school principal was asked to complete questionnaire material. Second, it seemed possible that the guidance officer in each school might be better informed than the principal on many of the guidance program details. Third, guidance and counseling was one part of the school's program on which it seemed desirable to gather more detail since

guidance people are those for whom Project Talent's research efforts would be likely to have the greatest meaning. And fourth, current information on guidance programs and on counselors in our high schools was deemed an appropriate topic for the study. The guidance field had been expanding quite rapidly suggesting that information gained from earlier studies would be out of date.

It seemed desirable to gather information on guidance programs for several reasons. We wanted to investigate on a national basis the ways in which high schools provide for guidance. Questions about the school's facilities for guidance, the number of counselors employed, administrative arrangements for guidance, the extent to which tests are used, and other topics seemed relevant. We wanted to learn about the ways in which guidance programs are integrated into school programs. In order to do this, we would have to have information that would enable us to compare the school characteristics with guidance program characteristics. An example of such a comparison might involve comparing school size with the number of counselors employed. With our data, we would be able to compare schools on their guidance programs as well as on other characteristics. Comparisons of this sort might prove useful to administrators in suggesting needed changes.

We wanted to correlate information on the guidance program with the student's test scores and questionnaire responses. We thought it desirable to be able to investigate, other factors equal, if guidance has a measurable correlation with students' test performance, aspirations, or plans for

college. We wanted, too, to correlate data on guidance programs and procedures with some criteria of student behavior including quitting high school, attending college, obtaining a scholarship or a student loan, or choosing an appropriate vocation. Research on the last two topics would help to clarify the role that guidance plays in the development of our high school students. For example, it might be expected that if guidance is being effectively carried out in our schools that the number of students who quit school might decrease, that deserving students would be helped to gain scholarships or student loans, and that students would choose colleges and vocations more appropriately.

The Project Talent Study of Guidance and Counseling

The Guidance Program Questionnaire and the Counselor's Questionnaire were developed in about the same way as the School Characteristics Questionnaire described in the previous chapter. The Counseling and Guidance Panel recommended topics and issues for study. The staff reviewed these recommendations in the light of current research and the objectives of the Project and submitted preliminary versions of the questionnaires to the panel for review and comment.

The first questionnaire submitted to the panel included questions about the guidance program and about the counselors in a single form. After reviewing this form, staff and panel agreed that the questions about the counselors should be placed on a separate form. In this way, each counselor could answer questions about his own education and experiences. The newly revised questionnaire was

submitted to the Panel for a final review at its meeting in October 1959. The suggestions and recommendations of the October panel meeting were then incorporated with the revised questionnaire to make the final form of the questionnaire.

Each school was asked to complete two questionnaires, the Guidance Program Questionnaire (GPQ) and the Counselor's Questionnaire (CQ). The GPQ was completed by the guidance officer or the principal, if the school had no guidance officer. The CQ was completed by each counselor in the school who devoted twenty percent or more of his time to guidance and counseling duties. The GPQ included 24 pages of questions and over 200 items of information. The CQ included six pages of questions and about 75 items of information. In most cases the questionnaires were pre-coded, but write-in responses were permitted for questions on which choices other than those presented might be possible. We felt that the variety of school practices were not necessarily completely known at this time. Therefore, well-documented choices were presented in pre-coded form, but the option to write in another response was presented. Answers were marked directly on the questionnaires. Upon return to the Project office, each questionnaire was reviewed for completeness. Incomplete questionnaires were returned to the school for completion. Next the questionnaires were coded and the data were punched on IBM cards for processing.

A variety of item types was used, including multiple choice, yes-no, rating scales, check list, write-in, and

paragraph. Following are some sample items:

4. How many years ago did your school organize a guidance program?

( ) 1. Less than 1 year ago (this school year)
( ) 2. 1 year ago
( ) 3. 2 years ago
( ) 4. 3 years ago
( ) 5. 4 years ago
( ) 6. 5 years ago
( ) 7. 6 or 7 years ago
( ) 8. 8 or 9 years ago
( ) 9. 10 to 15 years ago
( ) 10. 16 years ago or longer

11. For each of the following, how often does the counselor call the student's parents for a conference? Mark your answers as follows:

1. Always
2. Frequently
3. Occasionally
4. Rarely
5. Never

a. ________ For disciplinary reasons
b. ________ Low grades

The topics on which information was requested were as follows:

The Guidance program Questionnaire:

1. Scope of the program. This section asked whether or not the school had a guidance program. If there was a program in the school, further questions were asked about the size of the guidance staff, adequacy of facilities for guidance, and referral facilities in the community.

2. Types of aids and guidance provided by the counselors and the program. This is partly a matter of problems brought to counselors by the students and partly a matter of the counselor scheduling conferences with students and parents. It also includes special courses such as occupations, and means of imparting

educational and vocational information.

3. Past growth of the program. In this section questions are asked about expansion of the guidance program in recent years.
4. Plans for expansion. This section is similar to the previous one but asks about additions to be made to the guidance program in the next few years.
5. The testing program. This section is largely concerned with nationally standardized tests and inventories. There are many kinds of tests and inventories including the well-known intelligence tests, aptitude batteries, achievement batteries, interest inventories, adjustment inventories, etc. In addition, there are lesser known tests of special aptitudes such as those used to assess aptitude for art or music.

   Equally important are the express purposes for which tests are used. The same test may be used to provide information to students and parents on school progress, educational potential, or occupational potential. Similarly, the same test might be used to evaluate the progress of the school, to assign students to curricula, or to provide information to prospective employers, colleges, or scholarship agencies.

The Counselor's Questionnaire:

1. Duties performed. This section included questions on the proportion of time spent in guidance and counseling duties versus other

functions, such as teaching, administration, and the like. Information was obtained on the proportion of their guidance time spent in various duties, such as counseling, testing, staff discussions, preparing reports, etc.

2. Education and training. Questions included degrees obtained, coursework related to guidance and counseling, and how their training in guidance and counseling was obtained (i.e., through college, the school system, or on the job).
3. Experience. Questions were asked about the number of years of experience in teaching and in guidance and counseling.
4. Professional certification, participation, and interest. Professionalism in the guidance field may be reflected in a number of ways. Questions were asked about state or local certification, professional organizations of which they are a member, meetings of professional societies which they have attended, and professional magazines which they have read.
5. Objectives of guidance and counseling. This question explored the counselor's point of view on the important contributions he can make to the lives of his students.
6. Paragraphs. Two questions were included for which the counselor was asked to provide his views at length. One question asked the counselor to list the activities that should ideally be

included in a guidance program, and the amount of time to be devoted to each. The second question described a hypothetical "case study" and asked the counselor to tell how he would handle the situation.

7. Other questions asked about the counselor's sex, age, and salary.

The analysis of the data involves statistical computations from which useful information can be derived. Among the statistical analyses that will be performed are the following:

1. Weighted frequency distributions or "population estimates". A "population estimate" approximates the frequency distribution in the nation from the distribution in the sample. In Project Talent, we drew a 5 percent sample, therefore all of the population estimates will be based on this sample. Some examples of the information we plan to obtain from the Guidance Questionnaire and the Counseling Questionnaire using this method include estimates of the number of schools that have formal guidance programs, the number of full-time and part-time counselors in the country and the number of counselors with differing years of experience. The same type of analysis can also be made for regional areas and for urban-rural areas within the country.

2. Cross-tabulations. Questions can be cross-tabulated in order to gain more detailed information about guidance programs and counselors. For example, one analysis we anticipate conducting will compare the school enrollment with whether the school has a guidance program. From this comparison, we would be able to learn whether larger schools have guidance programs more frequently than smaller schools, or whether the size of the school makes any difference at all. In the Counselor's Questionnaire, we plan to compare such things as the counselor's age with his college coursework in guidance.
3. Proportions or percents. Among the proportions that will be computed are the proportion of the schools that have guidance programs, the proportion of counselors of each sex, the proportion of counselors with the master's degree, and the proportion of counselors in each region of the country. Many others will be included.
4. Correlation. We will calculate coefficients of correlation which will indicate the degree to which certain guidance or counselor characteristics are found together. For example, it will be possible to determine the extent to which large schools have guidance programs with many counselors. These calculations permit consideration of many

characteristics at once instead of individually. Results from studies of relationships are often more enlightening when the behaviors themselves are quite complex. Methodology will also involve use of multiple correlations and partial correlations.

With these analytical methods, we plan to study some selected aspects of guidance and counseling resources in American high schools. This study will provide data on how the nation stands at the time of testing on each of the characteristics of guidance programs and counselors mentioned in the outline of topics. This information should be useful to leaders in education and guidance for evaluating needs for counselors, for counselor training, and for guidance facilities. In addition, the information might be used by school administrators, principals, and guidance officers as an aid for evaluating their own schools' program of guidance by comparing themselves to similar schools.

Research plans also include investigations of some of the ways in which guidance and counseling may affect our students. For example, we will attempt to investigate the relationship of guidance to high school drop-out rate. Research will also be conducted on the relationship of guidance to helping students formulate college plans, obtain scholarships, or choose a vocation.

The results of Project Talent are expected to contribute greatly to the body of research information guidance counselor's use in helping students. The counselors first concern is, of course, to aid the student, using all

possible resources at his disposal. Among the tools needed for educational and vocational guidance are nationally standardized tests of aptitudes and abilities, and information on personality, interests and background that has demonstrated relevance to educational and vocational success. Project Talent will provide a great deal of this information through its research results. Although Project Talent tests will not be used commercially, findings of the study should contribute to the development of better methods for identifying each student's talents.

CHAPTER XI

THE MINDS OF AMERICAN YOUTH

Charles Martin

Left tackle Charles Martin (name fictitious) sat astride a chair in the classroom, his huge bulk filling it and overflowing. He was writing a five-minute theme for Project Talent on the topic "What High School Means to Me."

This is what he wrote:

> "If I were to be dishonest, I would say I obtained much from high school; but alas, to my dismay, my honesty provides that I should tell the truth. It meant a good time!"

Charles (age 18) is the son of a small businessman in comfortable circumstances. Both of his parents were high school graduates. He learned to dance at twelve, has two dates a week, and expects to marry at nineteen. He expects to go to college and study engineering. He has had six semesters of mathematics, including two months of calculus, and has made straight A's in his mathematics. His other grades are mostly A's. However, he has had no foreign language. Charles is interested in many things. He says he would like very much to be a Naval officer, chemical engineer,

aeronautical engineer, electronics technician, professional athlete, electrical engineer, mining engineer, sports umpire or referee, mechanical engineer, mathematician, Air Force Officer, college professor, psychologist, member of the President's cabinet, judge, U. S. Senator, U. S. Congressman. He expresses a strong liking for basketball, football, track, algebra, exploring, and baseball.

For his theme on "My Views of an Ideal Occupation", Charles wrote:

> "It seems only natural that one would crave an ideal occupation. To be perfectly frank, I haven't much conception of what I wish to enter. Perhaps as college progresses my solution will show itself."

Charles apparently is getting off to a fortunate start in life. He appears to have considerable talent, and plans to go on to college. However, a number of questions arise in connection with how one might consider Charles' case from the point of view of a counselor. For one thing, what of the lack of a foreign language? Will this stand in the way of his admission to a college that can make maximum development of his talents? Will he end up later in life regretting his lack of a foreign language?

Are Charles' work habits what they should be to continue to make good grades at a really tough engineering school? Is he being sufficiently challenged by the school he has been attending?

Charles is a good example of a student with high talents who makes good grades and plans to go to college and yet might have profited by the right sort of counseling in the earlier high school grades.

Sophia Martinotti

Sophia (age 17, name fictitious) wants to be a private secretary; she doesn't want to go to college. For her theme she wrote:

> "High school means the necessary education in order to get the job I want after high school. If I quit high school I probably wouldn't be able to be a private secretary as I hope to be. If I got a job without completing high school I wouldn't get as good a job as I want and I also wouldn't get enough pay."

For her other theme on an ideal occupation, she wrote:

> "After I graduate from high school, I would like to be a secretary (if possible, a private secretary). I would like to earn about $3,500 to $5,000 annually. I think a secretary earns enough money to live comfortably with a permanent job and stable income."

Sophia lives in a six-room apartment with her mother, father, and five brothers and sisters. Both parents work and are union members. She expects to marry at twenty-one or twenty-two years of age and expects to have four children. She expresses a high degree of interest in such occupations as airline hostess, secretary, typist, housewife, and also such activities as doing clerical work, helping the poor, learning about diseases, and swimming.

She had her first date at thirteen or fourteen, and dates about once a week. She has gone steady twice.

She reads over twenty-one books a year, but mostly books classified as adventures, histories, or love stories. She studies one to four hours per week and is taking a commercial course. She has had three or four semesters of math and six semesters of foreign language in the commercial curriculum. Her grades are B's or better except in science and foreign language, where they are C's. She says most of her friends don't expect to go to college, and that she

would not borrow money for college. Sophia wrote two well-written themes indicating a potential for profiting by college attendance. Her patterns of information and school grades indicate also that she might profit from instruction at an appropriate college.

Many of those qualified for college but not planning to attend are similar to the case of Sophia. In dealing with a case of this sort, the counselor needs the help of accurate scholastic ability measures in order to help predict how well she would do at various kinds of colleges. Counseling is also very important in these cases at the time the basic choice is made to take a commercial rather than an academic curriculum.

Sophia also doesn't believe that college has much of value to offer to her. What could the counselor and the school have done to give her a more favorable attitude towards college? What will Sophia herself think in future years of her decision not to go to college?

Fred Stone

Fred Stone (name fictitious) wrote:

> "I think a ideal occupation would be that of a Dentist because people will alway have teeth to be fists."

And this one:

> "High School means to me it is where you get your finl egcation be for going out in the world."

Here is how Fred describes himself, his family and his life plans:

Fred was 16 years old when he wrote these themes. He's about 5'10" tall, weighs about 155 pounds, and is in the 11th grade. He lives with his mother, father, two older sisters,

and one younger brother in a house which has 11 or 12 rooms. They are buying the house, which has a value between $10,000 and $15,000. They live in a community of about 25,000 in a western state. His father is about 52 years old, a semi-skilled worker for a rather large company. He is not a supervisor. His mother is 10 years younger. She has been a nurse, but is primarily a housewife now. Fred considers his family to be "well-to-do". The family income is between $9,000 and $12,000.

What is his family home like? Fred has a room of his own, his own desk and typewriter. The family has most of the modern conveniences--electric dishwasher, automatic washer and dryer, refrigerator, vacuum cleaner, food freezer, telephone, radio, phonograph, sterling silver, paintings, wall-to-wall carpeting, and drapes. They have hi-fi, records, and quite a bit of sports equipment. They have many hand and power tools.

Although they have between 100 and 250 books in their home, the main reading apparently consists of magazines--ladies', men's, movie, mechanical, news, business, etc.

The family has two cars, the newest of which is a 1954 model.

Fred's parents are both native-born. His father is active in church, but his mother is not. They both do some work in the PTA. His mother is active in the League of Women Voters and other civic organizations, but his father takes no interest in politics and does not belong to any civic groups. Fred doesn't know how much education his folks have.

Fred says he is in excellent health. He had the usual

THE ORIGINAL DOCUMENT WAS OF POOR
QUALITY. BEST POSSIBLE REPRODUCTION
FROM COPY FURNISHED ASTIA.

childhood diseases, but has only been out of school a day or two at a time. He sleeps 8 hours a night. On weekends he stays up until 12. Although he learned to drive at 13 (driver training school) he isn't allowed to take either of the family cars.

He first began dating at 13--has about 3 dates a week now; he has gone steady once.

Fred says he really isn't interested in school. This, he says, is why his attention wanders; he can't concentrate; he doesn't keep up to date on assignments. In fact, he's behind most of the time. He doesn't feel that the courses will help him after he gets out of school.

When he gets an assignment, he isn't always sure of of what he is to do. He's not a fast reader (he pronounces words to himself), but he does understand what he reads. When he turns something in, most of the time his teachers tell him that it's sloppy. He's careless about spelling and English, which pulls his grades down. Not paying attention in class has hurt his grades most of the time, too. He doesn't consider school work a challenge.

He studies on an average of seven hours a week. He reports that he is a C to D student. His school record, themes, and very poor scores on all the information tests agree in indicating a very poor prognosis for success at college.

Fred has been active in clubs. He played on two athletic teams, and reports he was captain of one. He learned to dance when he was 13.

Fred likes to hunt and fish, and has often gone in for gardening, sports, and repairing cars. In addition, he goes

to concerts, plays, museums, etc.

He helps out around the house about 8 hours a week, and works 21 hours a week to earn money. If he needs more spending money, he gets it from his family.

In the last year, he has read fewer than 5 books outside of school requirements, and these are of the western adventure, science fiction, comic book types.

Fred made up his mind in the 7th grade that he wanted to be an engineer. He has had many discussions with his teachers and counselors regarding college. Perhaps as a result, he has decided that he will likely to go to a business school--part-time--or a junior college. His plans are indefinite--he <u>may</u> go, but expects it to last less than four years. His parents want him to have a college degree, but he says he can expect no financial help from them. Rather, he will have to depend on other sources--working, scholarships, etc. He would not consider taking out a loan to get through school.

Despite his lowered level of aspiration, he expects to be earning $25,000 twenty years after high school graduation, and wouldn't be satisfied with less.

He expects to go into service by getting an ROTC commission. He dislikes the thought of a permanent military career, however, but might consider it if the salary were better.

Fred considers job factors in the following order of importance:

Security
Freedom of decision
Good income

Friendly associates

Promotion

Importance of work

However, to improve himself from one job to another, he puts the factors in this order:

Good income

Promotion

Better Supervisor

Better associates

Importance of work

Security

He hopes and expects to be "well-to-do", and have one child.

Fred presents a tough problem for a counselor. Although he has always wanted to be an engineer, he appears to lack the qualities which will move him toward that goal. If he could even get into college with his grades, his habits, his lack of interest in school, his low test scores, and his past history indicate that he is not a good bet to succeed. Even after dropping back to a more realistic goal--business school or junior college--he remains uncertain and indefinite about his plans......Maybe he'll go.... but he expects it to be for less than four years.

Fred may be typical of many students his age who expect to succeed almost by a magical process, without knowing what it really takes to become a $25,000-a-year person. His counselor should begin now to reorient his goals to a more realistic level.

Perhaps Fred has subconsciously done this to some degree when he says that:

> "High School means to me it is where you
> get your finl egcation be for going out
> in the world."

If Fred's test scores show up some specific aptitudes outside of the academic area, he might still be able to be guided into a useful and important occupational niche. Right now his potential in life does not look encouraging without proper guidance.

The two themes written by each student reveal knowledge of effective expression, punctuation, spelling, and English usage, but also seem to reveal in most cases something about basic attitudes, values, and other important insights regarding the student's thinking and his basic personality and temperament.

Initially, it will not be possible to score both themes for all 440,000 students, but it is planned to score a small sample of them to provide a base line for comparison. Later, during the follow-up studies, it is planned to study the themes of persons who end up in groups or categories of special interest, such as statesmen, scientists, naval officers, and the like, or also groups who are failures in life in spite of having many talents.

Experience with the themes in a pilot study indicates that it should be possible to derive from these themes measures of altruism, scholarship, basic value systems, and the self-concept. This may result in some contribution to the basic arts of personality, attitude, and values measurement. It should also result in substantial new insights into the structure of motivation for specific careers, including military careers.

Initial analyses of the themes should determine what kinds of scores can be obtained from them. The scores

derived from the themes would then be related to all of the other measures available on the same individual in order to help define what the new scores mean in terms of consistency and overlap with other known measures.

The first step in the analysis of the Talent results will be a complete analysis of a carefully drawn sub-sample of all the students who took the Battery. This sub-sample will be used to make a thematic analysis to tabulate the frequency of occurrence of specific theme topics. In this thematic analysis a "theme topic" is defined as being an identifiable basic attitude value or the like, such as altruism, social activities, learning something, making money, military service, or the like. After this has been done, a limited number of the more common "theme topics" will be integrated into a scoring system.

While various scoring systems should be experimented with, it is believed, on the basis of pilot studies, that the most promising approach is to have trained readers read the theme, make a number of ratings, and at the same time punch these ratings into a card. One score might be obtained by rating the theme on a 9-point scale in terms of adequacy of writing. The other ratings would be mostly on a presence-absence basis where the rating indicates the mention of a given predefined topic. These readings and ratings would be obtained independently and checks would be made on the consistency of scoring. After the preliminary set of scores had been derived. from the themes, they will be subjected to statistical analysis to determine their interrelationships as well as their relationships with the other measures available on the same subjects.

The research on the themes should produce a valuable new approach to the scoring and quantitative analysis of samples of writing for both quality of writing and for many measures of value systems, attitudes, self-concepts and temperaments. The final outcome should be a generalized system applicable over a wide range of topics and content of compositions.

While the final scoring system for the essays has not yet been worked out, it is believed that the examination of a few samples of these can illustrate some of the different types of values which seem to emerge.

Job Training

Many students see education essentially as specific job training. This is true both at the college and the high school terminal level. Some samples of these are given below:

> Girl aged 15. (Occupation)
>
> "I would like to go into the field of costume design after a few years or so of college. Before I enter college (if I decide to go) I would like to go to business school for a few months. So in case I don't get a job in the field of costume design right away, I will be equipped for secretarial work."
>
> Boy aged 17. (High School)
>
> "In everyday life I think that school doesn't mean much to me. I think that I am going because my mother and father are keeping me there. But when you get down to brass tacks, school is the basis of my entire future as well as every other student's in my class. Without going to high school, it would be very difficult to get a good job which would satisfy you. So I think that every student should keep on going to school until he graduates and then he would have a better chance to get a job."

Social Service

On the other hand, a number of students have very strongly expressed desires to be of social service to others. Examples of this are given below:

> Girl aged 18. (Occupation)
>
> "To me teaching in the elementary school level is the ideal profession. It covers an excellent field and requires dedication. This ideal occupation should and does hold my interest. What could be more rewarding than molding and shaping the minds of tomorrow's citizens. Money cannot measure the satisfaction gained from this profession."
>
> Boy aged 19. (Occupation)
>
> "The ideal piece of work I would enjoy having would be something in line of social working among delinquents in New York or maybe a job as an artist. These jobs are both creative jobs and I think more suited for this type of work, for it gives me a chance to express my self and maybe helping a job or two like myself."

Money

Of course, a common theme in many essays is money or material goods or material success, expressed in a number of ways. Examples are given below:

> Boy aged 17. (Occupation)
>
> "I believe that the ideal occupation is a lawyer. A lawyer doesn't have to get paid by the week or month, but by the case. Some cases yield him more than the average persons total wages for a year. But of course there are all different types of lawyers, such as criminal, Business--"
>
> Boy aged 17. (Occupation)
>
> "a millionaire"

Security

Some students are very security conscious and this seems to loom large in their thinking and planning. Examples are given below:

Girl aged 15. (High School)

> "In my mind high school means a knowledge of things I can carry with me into the years. It sets forth a lot of memories I can cherish and pass down to my children. But mostly high school means security for the future."

Boy aged 16 (Occupation)

> "I would like a good occupation with security and room for advancement. I would like a job with a salary of $15,000 or more a year and good working conditions and nice people to work with."

Learning

Some students seem to give considerable importance to learning as a goal in itself. Some give much importance to achievement, the attainment of fame, becoming successful.

Girl aged 15 (Occupation)

> "To me an ideal occupation is teaching. Teaching gives you a satisfaction in knowing that you are helping others to learn. Teaching also broadens your own personal knowledge. I want to be a teacher because I think that teachers are among the most called for professions of this day. Some students are not gaining the knowledge they should because the teachers are not teaching properly. I plan to teach to the best of my ability and help others all I can. This is my idea of an ideal occupation."

Girl aged 17. (High School)

> "I have always enjoyed school and done my best to get good marks. High school is the last lap in education before college and extremely important. High school is the educational foundation of our education to come. It is a determining factor on what our future lives shall be like. High school also helps to develop one's personality and ability to get along with others. This too is part of one's education."

Social Activity

Other students stress social activity.

Boy aged 17. (High School)

> "High school is a means to prepare myself for college not only scholarly but socially. Although I expect high school to prepare me as much as possible, for the rigors of college classes, I also expect it to prepare me socially. A fine mind is very good, but even a genius may become a nervous wreck."

Girl aged 17. (Occupation)

> "I would like to be in a profession where I could meet all types of people. Mostly I want to be rather commercial sort of professional. I would very much like to be a certified public accountant secretary or office clerk."

It is hoped that the information obtainable from the essays will help supplement the objective test and questionnaire information in giving a more complete description of the youth in the study.

CHAPTER XII

WHAT HAVE WE LEARNED FROM PROJECT TALENT?

First and foremost, we have learned that a national census of the aptitudes and achievements of American youth is feasible and can be done. The data-gathering phase has been completed successfully with excellent cooperation from all the individuals involved. Preliminary inspection and analysis of the data indicate that the students cooperated and were well motivated, and the teachers in the schools did a good job of giving the tests and collecting the data.

While this in itself may not appear very difficult to accomplish, consider that nearly one and one-quarter million individuals had to cooperate to make the study possible. Approximately 5000 school board members and 750 superintendents had to agree to cooperate and to devote two full days of their instructional program to the collection of data for the Project. Approximately 750,000 parents and guardians and approximately 440,000 students were involved. Approximately 1400 principals and headmasters, 1500 counselors, and 18,000 teachers were responsible for the actual giving

of the tests and collection of data. Each school also filled out a 50-page questionnaire.

In most cases the students were tested by their teachers in small groups of 25 or so. The 90 Regional Coordinators took the Project to the schools and were responsible for all the field work. Approximately 60 advisers and consultants supported the staff of seven professional research personnel who were responsible for carrying out the Project. Probably no research project in history has involved so many individuals so closely, and certainly no large research project of this magnitude has been as thoroughly decentralized. Non-cooperation by any link in this chain would have made it impossible to complete the Project.

In particular, the schools and teachers really outdid themselves in carrying out such a massive testing program. Many schools which had never before done very much testing carried out the program with enthusiasm. Under the general guidance of the Regional Coordinators, the teachers did a very satisfactory job of giving the tests and collecting data.

The enthusiastic response and cooperation of this large group indicated that, collectively, they have great faith in the vitality and adaptiveness of the American systems of education. While these hundreds of thousands of individuals differ in many important ways in their approach to life and education, they united in support of Project Talent in the belief that it is good to learn more about our youth and our schools and that this knowledge will help us to adapt to the challenges of our modern world.

Now that the data have been collected, what will we learn about our youth? From the preliminary analysis of the data during the first year or two, we will learn a great deal about the basic talents, achievements, aims, perceptions, aspirations, plans, and values of American youth.

We will learn what youth wants to be. We will know what proportion of boys and girls of various ages in different grades would like to become chemists, auto mechanics, musicians, research scientists, actors,or actresses, professional athletes, doctors, lawyers, politicians, high school teachers, college professors, religious workers, President of the United States, and so on.

We will learn what youth wants to do. We will learn what proportion of the various categories of youth like such activities as reading, fishing, basketball, saving money, buying stocks, studying algebra, becoming a millionaire, hunting, reading poetry, playing baseball, campaigning for political office, owning their own business, helping their parents, visiting art galleries, playing chess, watching television, studying foreign languages, teaching children, reading literature, going to school, military drilling, and doing many other similar things.

We will learn what youth knows. We will learn what young people know about science, guns, fishing, military activities, architecture, medicine, animals, games, languages, business, foods, boats, photography, theater, geography, art, swimming, literature, the Bible, music, social studies, farming, mathematics, electronics, mechanics, etc.

We will learn what youth will be able to do. From the aptitude and achievement test scores we will be able to estimate the capability of our youth for such important activities as attending different kinds of colleges, and for absorbing training in many important activities in military and civilian life such as electronics, mechanics, clerical and secretarial work, or learning foreign languages.

We will sketch out the life history of American youth. For various categories of our youth we will have information on the percent engaging in numerous hobbies and activities. We will learn about their work experiences--their part-time work and summer jobs and the number of hours they spend per week doing chores at home. In the very important area of dating, we will learn how old they were when they had their first date, how many times they have gone steady, how often they have dates, and also at what age they learned to dance. This information should be extremely valuable to study as a factor influencing study habits, staying in school, going to college, making grades, and the like. We will learn much about the reading and study habits of young people in relation to the grades they make and the way they carry forward their education. We will learn about their marriage plans. This is very important in regard to whether they go to college. We will learn much about the driving and automobile activities of youth. It is well-known from previous studies that these are very important in relation to studying and making good grades. We will learn a great deal about the future plans of American youth in relation to college, military careers, planned occupations, professions, etc.

From all of this will emerge comprehensive pictures of the various types of American youth in relation to their schools and homes. This information should be of great interest and value to all those concerned with educating our youth and dealing with their problems.

Studying Effects of Teaching

There is much discussion these days about changes in American education. Proposals have been made to get rid of all the small schools, make the classes smaller, pay the teachers more, use different instructional methods, revise the curriculum, and many others. All of these cost money, and probably most of them would have a favorable effect on the schools. However, very little is known about what the relationship is between class size and how much the students learn; or about the proportion going to college or the like, with other things held constant, such as the kind of community the school serves. Likewise, we have little information on the effects of any school practice on the ultimate goal toward which the schools are working. In particular, very little is known about how different ways of educating the gifted and those with special talents affect the extent to which they obtain maximum utilization and fulfillment of their talents.

To learn the answers to such questions, it is necessary to study a very large number of students and their schools and to do a thorough nation-wide follow-up of the students to see to what extent their later behavior coincides with the planned desirable outcome goals of the school. Within the first year follow-up of Project Talent we will begin studying

4. Why do some able students dislike studying, and how can the schools best encourage them to study more?

5. When and how do students make important decisions regarding their educational and vocational careers? How may the teachers, counselors, and parents best assist youth in making these decisions more wisely?

Future Publications

As the data from the testing and follow-up studies are analyzed they will be reported in later volumes of this same series. Volume II, _Educational and Guidance Programs in the American High School_, and Volume III, _The American High School Student_, will be published in about twelve months. Volume IV, _The American High School Graduate_, should appear about a year later.

The general planned content for Volumes II and III is given below:

Volume II

Educational and Guidance Programs in the American High School

A. Introduction: Brief over-all view of Project Talent

B. Collecting the information

   1. The tests--brief description

   2. The questionnaires

      a. How developed

      b. How administered

C. Life history of the American school

   1. Types of secondary schools

   2. The community served by the school

   3. Expenditures

   4. The Principal

   5. The Faculty

6. Class size
7. Admission, assignment, promotion, and grading policies
8. Homework
9. The school building
10. The library
11. Double shifts
12. Summer school
13. Classes for special groups
14. Treatment of the gifted
15. Foreign language instruction
16. Religious education
17. Absenteeism
18. Student turn-over
19. Drop-out rate
20. College attendance rate
21. Extra-curricular activities
22. Courses offered

D. The guidance programs of the American school

1. Number of persons assigned full-time
2. Facilities for counseling
3. Facilities for referral
4. Problems treated by counseling
5. Counseling procedures
6. Conferences with parents
7. What counselors do
8. Policy on reporting test results to students
9. Recent changes in program
10. Plans for expansion of counseling
11. Plans for expansion of testing

12. The use of tests
    a. What kind
    b. How used
13. Vocation information programs
14. The counselors
    a. Sex, age, salary
    b. Percent of time in counseling
    c. Duties assigned
    d. Distribution of time
    e. Training and experience
    f. Approach to counseling

E. The results of educational and guidance practices

1. How to determine results of school practices as they affect life outcomes
2. Effects on drop-out rate
3. Effects on college attendance
4. Effects on occupational plans
5. Effects on life adjustment
6. Effects on amount learned
7. Effects on attitude toward learning
8. Effects on reading skill
9. Effects on writing skill
10. Effects on spelling ability
11. Effects on motivation for science careers
12. Effects on motivation for mathematics careers
13. Effects on economic attitudes
14. Effects on military career plans

F. Plans for further volumes

Volume III

<u>The American High School Student</u>

A. Introduction

B. The tests--what they measure

  1. Description of tests
  2. Means, standard deviations, and r's--treated non-technically
  3. Relationships to other common tests

C. The talents of American youth

  1. Test scores of boys and girls
  2. Students with high or low grades in key subjects
  3. Students grouped by grade, age, father's occupation, student's planned occupation
  4. Country vs. city, regional differences
  5. Large vs. small schools
  6. Plans for college
  7. Plans for military service
  8. Drop-out intentions
  9. Marriage plans
  10. Science career plans

D. The life history of American youth

  1. Distributions by grade and sex for items on:
     a. Hobbies
     b. Clubs
     c. Sports
     d. Work
     e. Spending money and allowances
     f. Dating, dancing, going steady
     g. Reading habits
     h. Study habits
     i. Pupils' prior school history
        (1) Changing schools
        (2) Absenteeism

    - (3) Promotion
    - (4) Courses taken
    - (5) Grades received
    - (6) Counseling received

- j. Family and home
    - (1) Father--age, occupation, education, military service, whether born in U.S., language spoken, organizations belonged to, income
    - (2) Mother--age, education, working status, income
    - (3) Brothers and sisters--twins, number of children, birth order, education
    - (4) Housing--rent paid or value of home, number of rooms in home, number living in home, books in home, tools, magazines, appliances
    - (5) Where family lived previously
- k. Age expect to marry, number of children expected
- l. Health status--height and weight, vision, hearing handicaps, disease history, amount of sleep
- m. Special awards received
- n. Driving--age learned to drive, car ownership, driver training received at school
- o. Plans
    - (1) Continue high school
    - (2) Go to college--what kind, when, how many colleges applied to, how far from home, where expect to live, taking National Merit Scholarship Tests, reasons for college attendance, attitude toward college, college plans of friends, attitude of parents toward college
    - (3) Military service--lifetime career intentions, choice of service, reasons for plans
    - (4) Occupations--desired occupation, expected occupation, firmness of choice, when decision made, who helped make

decision, reasons for choice

(5) Financial plans--salary expected, savings account, stocks and bonds, real estate, life insurance, installment buying, reasons for saving, how much wealth expected

p. Cost of college

(1) Willingness to borrow under various conditions

(2) Amount of money needed

(3) Sources of support expected

E. What American youth knows

1. What high school students know about the following (distributions of sub-scores or typical information items in each area):

Bible
Words
Books
Music
History
Mathematics

Science
Space
Electronics
Mechanics
Farming
Home Economics

Sports
Guns
Business
Geography
Foods
Animals

F. How American youth reads

1. The Domain Test concept (sampling a defined domain or area of skills)

2. How many words high school students can use

3. How well they can spell

4. How well they can read

5. The correlates of reading skill--father's occupation, mother's education, high school grades, college plans, the characteristics of good vs. poor readers

G. The interests and activities of American youth

1. Science

2. Computation

3. Mechanical--technical

4. Skilled trades

5. Literary--linguistic

6. Social service

7. Public service

8. Musician

9. Artistic

10. Business Management

11. Sales

12. Office work

  13. Labor
  14. Farm and ranch
  15. Hunting and fishing
  16. Sports

H. The personal qualities of American youth

  1. Sociability
  2. Social sensitivity
  3. Impulsiveness
  4. Vigor
  5. Calmness
  6. Tidiness
  7. Culture
  8. Leadership
  9. Self-confidence
  10. Maturity of personality
  11. Conventionality
  12. Theoreticality
  13. Group-centeredness

I. Following the lives of American youth

  1. Follow-up plans
  2. What we expect to learn about:
      a. College attendance and success
      b. High school drop-out
      c. Entry into occupations
      d. Occupational satisfaction and success
      e. Military service
      f. Marriage, divorce, children
      g. Social behavior problems
      h. Psychiatric problems
      i. Creativity
      j. Migration patterns of students with different talents

Volume IV, The American High School Graduate

Plans are less definite for the exact detailed content of Volume IV, but in general, this will report the result of the first of the one year follow-up studies. The seniors who were tested in 1960, will be followed up in 1961. The material that we collect from them will then be related to all of the material collected on the students at the time they were tested in 1960 and also to the information collected about the schools at that time. A large amount of information should be reported in this volume regarding the relationship of the originally-gathered data to such follow-up measures as school attendance and success, the kind of course of study entered, high school drop-out, entry into occupations, occupational satisfaction and success, military service, marriage, social behavior problems, psychiatric problems, creativity, and migration patterns of students with different talents.

From all of this will emerge comprehensive pictures of the various types of American youth in relation to their schools and homes and their plans and aspirations. This information should be of great interest and value to all those concerned with educating our youth and dealing with their problems.